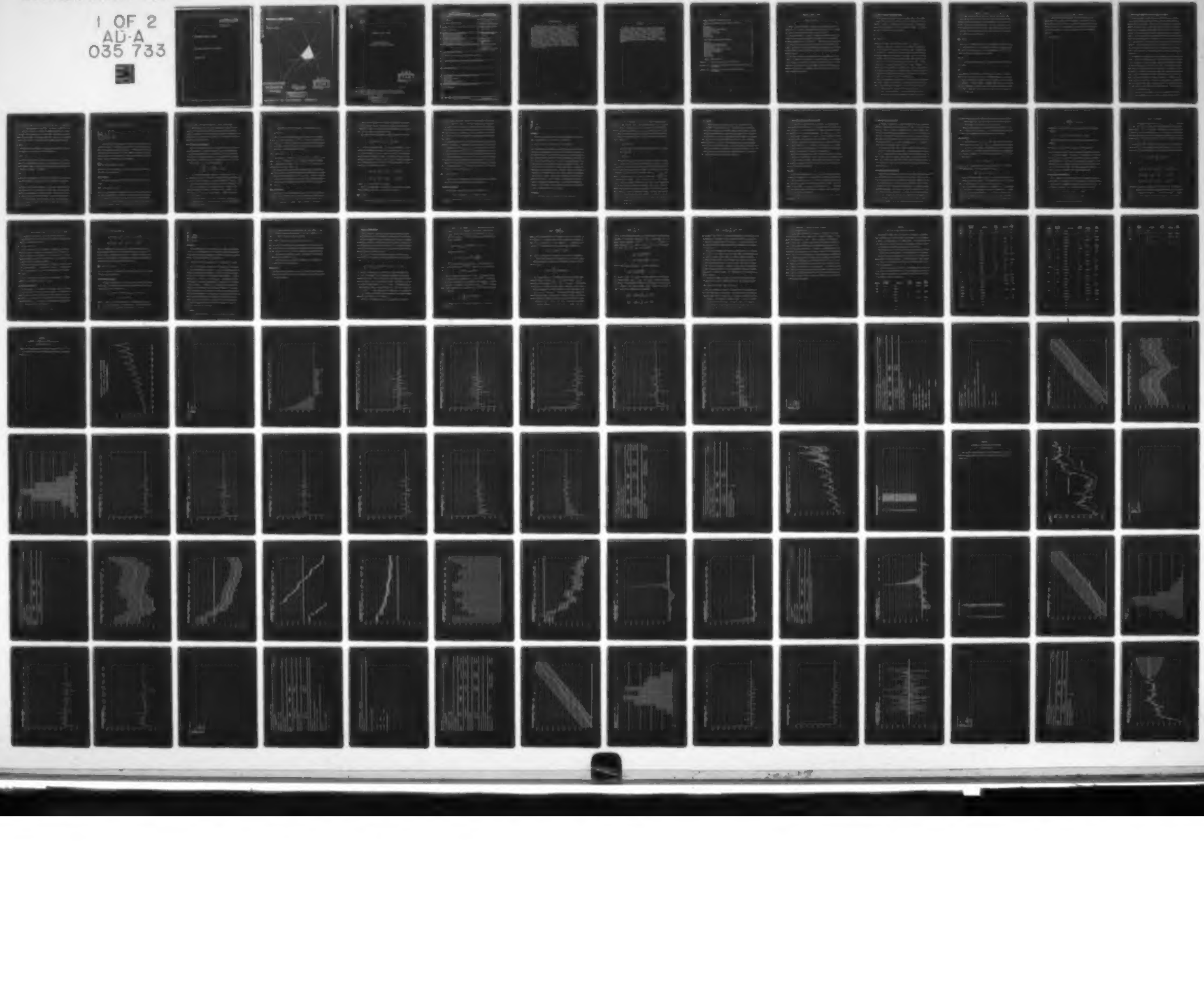
AD-A035 733 CALIFORNIA UNIV BERKELEY OPERATIONS RESEARCH CENTER F/G 12/1
EVERYMAN'S GUIDE TO TIMES.(U)
JAN 77 R R WILLIE AF-AFOSR-3179-77
UNCLASSIFIED ORC-77-2 NL

1 OF 2
AD-A
035 733



U.S. DEPARTMENT OF COMMERCE
National Technical Information Service

AD-A035 733

EVERYMAN'S GUIDE TO TIMES

CALIFORNIA UNIVERSITY, BERKELEY

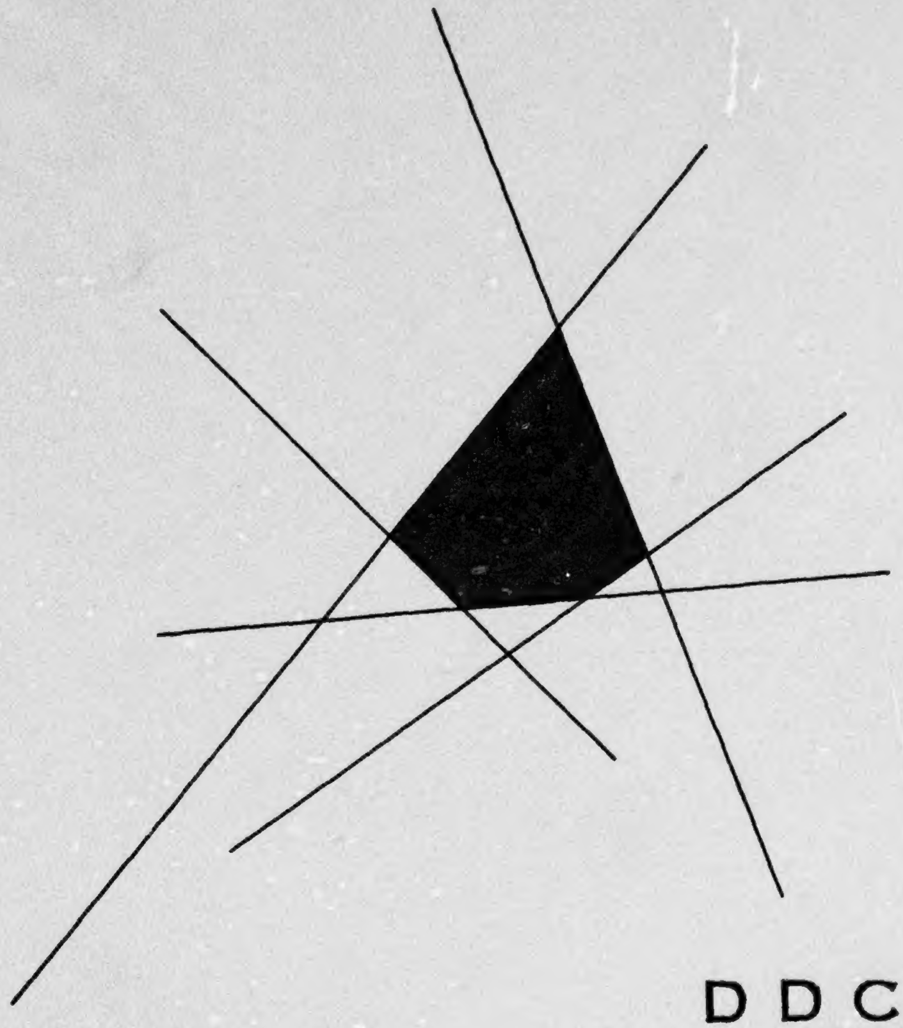
JANUARY 1977

ORC 77-2
JANUARY 1977

EVERYMAN'S GUIDE TO TIMES

by

RANDALL R. WILLIE



OPERATIONS
RESEARCH
CENTER

UNIVERSITY OF CALIFORNIA · BERKELEY

EVERYMAN'S GUIDE TO TIMES

by

Randall R. Willie
Operations Research Center
University of California, Berkeley

JANUARY 1977 ORC 77-2

This research was supported by the Air Force Office of Scientific Research (AFSC), USAF, under Grant AFOSR-77-3179 with the University of California. Reproduction in whole or in part is permitted for any purpose of the United States Government.

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER
ORC 77-2

2. GOVT ACCESSION NO.

3. RECIPIENT'S CATALOG NUMBER

4. TITLE (and Subtitle)
EVERYMAN'S GUIDE TO TIMES

5. TYPE OF REPORT & PERIOD COVERED
Research Report

6. PERFORMING ORG. REPORT NUMBER

7. AUTHOR(s)
Randall R. Willie

8. CONTRACT OR GRANT NUMBER(s)
AFOSR-77-3179

9. PERFORMING ORGANIZATION NAME AND ADDRESS
Operations Research Center
University of California
Berkeley, California 94720

10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS
2304/A5

11. CONTROLLING OFFICE NAME AND ADDRESS
United States Air Force
Air Force Office of Scientific Research
Bolling AFB, D.C. 20332

12. REPORT DATE
January 1977

13. NUMBER OF PAGES
99

14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office)

15. SECURITY CLASS. (of this report)
Unclassified

15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)
Time Series
Box-Jenkins' Methods
Forecasting
Transfer Function Modelling
Spectral Estimation

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

(SEE ABSTRACT)

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

ACKNOWLEDGEMENTS

More than half of the computer code for the TIMES program was written at the Ohio State University. A listing of the Ohio State routines was initially forwarded to personnel at the Operations Research Center, U.C. Berkeley, by Professor Nozer Singpurwalla of the School of Engineering Science, George Washington University. The comprehensive documentation of these routines by David J. Pack was a valuable aid in preparing the current version of TIMES. Professors Robert Oliver and Richard Barlow of the Department of Industrial Engineering and Operations Research, U.C. Berkeley, provided motivation for this project and a number of important suggestions. Professor David Brillinger, also of U.C. Berkeley, contributed the Fast Fourier Transform subroutine and most of the methodology behind the spectral analysis package, which is patterned after several of his own programs.

ABSTRACT

This is a user's manual for the TIMES computer program, an integrated system of FORTRAN subroutines for modelling of a single time series and for transfer function modelling with a single input series. TIMES is an extension and adaptation of a widely used system available from the Ohio State University Computer Center Library. This document is a refinement of an earlier guide prepared for classroom use during Spring Quarter, 1976, at the Department of Industrial Engineering and Operations Research, University of California, Berkeley.

TABLE OF CONTENTS

EVERYMAN'S GUIDE TO TIMES

by

Randall R. Willie

TIMES is a system of computer routines for time series analysis and forecasting. It is a tool for investigating both the structure of a single time series and the relationship between a pair of series.

The system is based on a library of FORTRAN subroutines acquired from the Ohio State University and designed for analysis of time series data utilizing techniques of Box and Jenkins [1]. This library has been extended at U.C. Berkeley to include a spectral package and a user-oriented control program. In addition, the original subroutine for plotting on a line printer has been replaced by more sophisticated routines that enhance and simplify graphical output.

The control program accepts time series data and a number of user instruction statements, each confined to a single 80-column card. The purpose of this writeup is to describe the control program input options. The user should have some familiarity with Box-Jenkins' methods and terminology, as presented in Reference [1]. Examples illustrating program output are provided in Appendices II and III.

I. GENERAL INSTRUCTION STATEMENT FORMAT

The instruction statements all have a common format. Card columns 1 through 5 contain a *keyword*, which is a specific string of five or less characters. Columns 7 through 80 contain an *entry list*, which consists of numeric character strings, or *entries*, separated by commas. For instance, this card is a valid instruction statement:

```
Col.
1    7
↓    ↓
IDEN 101,0,51
```

The keyword is 'IDEN' and the entries are '101', '0', and '51'.

Each entry represents a numeric value, so these strings may consist only of the characters 0 through 9, plus or minus signs, a decimal point, and the letter E. In general, any numeric character string acceptable in the FORTRAN language will be properly interpreted by the control program. For instance, entries that may be used are integer representations, such as '38.2' or '-.01', or scientific notation, such as '-1.0E-5' '.1E-01', or simply 'E25'. Improper strings like '1.0.0', 'FRED', '+1.0+E01', '.E01', '1.0E.1', and '+1E' will result in a SYNTAX error. Up to eight blanks may appear immediately preceding or following an entry string in the entry list, but blanks may not appear within the string; for example, the entries '10 45', '+ 1.0E5', or '1.0 E5' will result in a SYNTAX error.

The value represented by each entry must lie between an upper and lower bound. These bounds vary depending on the particular keyword and entry; a complete list is provided in Appendix I. If the value lies outside of the bounds, a BOUND error will be detected by the control program.

Each entry value is assigned to a particular parameter in the control program, and the control program determines the disposition of an entry solely by the number of commas preceding it in the entry list. For convenience, many of the entries are *optional*; that is, the user may choose to leave such entries out of the list, and the control program will automatically assign a default value to the parameters associated with these entries. However, if a particular entry is left out, unless all of the succeeding entries are also left out, the comma following that entry must be retained. On the first card below, the first entry has been left out; on the second card, the second entry:

```
IDEN  ,0,51
IDEN  101,,51
```

If there are n possible entries, and the user wishes to leave out the last n - k (optional) entries, say, the portion of the entry list after the k^{th} entry may be left out entirely. For the above example, if the last three entries are ignored, the following card suffices:

```
IDEN  101
```

In fact, if all entries are optional, an instruction card need only contain the keyword:

```
IDEN
```

Unfortunately, not all entries for every keyword are optional. *Required* entries must be provided by the user, or a REQUIRED ENTRY error will result.

Output from the control program includes a listing of the user instruction statements. If SYNTAX, BOUND, or REQUIRED ENTRY errors are encountered, appropriate messages will appear below the instruction statements containing these errors. Execution of the time series routines will not begin if any of these errors occur.

For discussion of particular instruction statements, it will be convenient to have a basic format notation. Individual entries will be designated by names in small letters; hopefully, the name will recall to the user the function of the entry, so a times series analysis may be carried out only with the aid of Appendix I. Names of optional entries will be enclosed in brackets. Thus, for instance, an instruction statement format is:

```
AR    [initv],ord,ifac
```

II. MODELLlNG AND FORECASTING WITH A SINGLE TIME SERIES

There are a number of approaches to extrapolation of a time series, but the TIMES program is primarily suited to one very general approach suggested by Box and Jenkins. Based on the first and second order sample moments of a given series, a particular model is chosen from the class of autoregressive integrated moving average (ARIMA) processes; the model parameters are then estimated by nonlinear least squares. The statistical adequacy of this fitted model in describing the data may be judged by techniques discussed in Chapter 8 of Reference [1]. At this point in the analysis, it may be necessary to modify the form of the model and reestimate the parameters. When the analyst feels that a satisfactory "fit" to the data has been obtained, the model can be used for forecasting. The principal stages in this approach are *model identification, parameter estimation,* and *forecasting.*

The TIMES control program permits some flexibility in performing these stages. The control program processes instruction statements and time series data in two phases. In the first phase, instruction statements are decoded and entry values and time series data are stored. The second phase begins when either the control program input stream ends or a RERUN instruction is encountered. If instruction statement errors have been detected in the first phase, execution terminates; otherwise, some additional error checks are made to insure that the input is logically consistent. Finally, the time series routines are executed. If the input stream contains an IDEN instruction, the identification stage is performed first. The estimation stage is performed next if an EST instruction has been provided. The FCST instruction controls the forecasting stage.

One execution of both phases of the control program will be called a *run*. Unless the RERUN instruction is used, a program run will thus be the same as a single computer job, and any or all of the stages of identification, estimation, or forecasting may be performed. The order of instruction statements in the input stream for a given run is completely arbitrary. Instructions having certain keywords may not be used more than once for a run; for example, multiple usage of the keywords IDEN, EST, and FCST is not permitted. Appendix I specifies for each keyword whether multiple usage is allowed.

Data Input

The DATA instruction prepares the control program to accept data values for the time series. The format of this instruction is

DATA npts

The entry *npts* gives the number of data points in the series. The DATA instruction must be immediately followed by a card describing the data format. A standard FORTRAN format item enclosed in parentheses should appear on this card, beginning in column 1. An example is

(8F10.0)

The format item will usually be of the form aFb.0, where b denotes the width in character positions of a data field, and a is the number of fields on a card.

The data cards follow the format card. Values are placed successively in data fields from left to right across a card; as many cards may be used as necessary to accommodate all time series points. However, all cards but the last must have a data value in each field. This value should contain a decimal point and may appear anywhere within the field, unless E-notation is used (such as 1.0E-02). In this latter case, the data item must be right-justified in its field.

A suitable data package would be as follows:

```
Col.
1    6    11   16   21
↓    ↓    │    │    │
DATA  12  │    │    │
(5F5.0)   ↓    ↓    ↓
3.4  12.2 0.   1.E-520.0
 5.2 3.4  9.8   6.    1.0
12.0 13.0
```

A project name may be assigned to the input series through the use of a NAME instruction. This project name is used as a title for graphs and data listings produced by computer routines. The NAME instruction has no entries, and consists simply of the string 'NAME' in columns 1 through 4. This instruction must be followed by a single card containing the project name in columns 1 through 80, as in this example:

```
NAME
WHOLESALE CUMQUAT PRICES 1930-1970
```

The NAME instruction is optional; if it is not used, the name is taken to be blank on all program output.

Identification

The IDEN instruction controls the identification stage and has the format:

```
IDEN  [nacor], [dplot], [nchi]
```

The entry *nacor* specifies the number of lags for which the autocorrelation and partial autocorrelation values will be calculated. These statistics are calculated for the input series and the first and second differences of the series. If *dplot* is nonzero, a plot of the series will be provided. Finally, *nchi* gives the number of autocorrelation values to be used in calculation of the chi-square statistic to test for white noise.

A single program run for the model identification phase may be made with only a DATA instruction and data package and an IDEN instruction. The user may also provide an SDIF or TRAN card if seasonal differencing or transformation of the input series is desired prior to identification. These instructions are described below.

Differencing and Transformation

Very often the input series will not appear stationary or will contain strong periodicities. The user may choose to apply a transformation or several difference operations to the input series to yield a stationary series that can be suitably modelled as an autoregressive moving average (ARMA) process. More specifically, the stationary series x_t is related to the input series X_t in the following manner:

$$x_t = \left[\prod_{i=1}^{n_s} \left(1 - B^{o_i}\right)^{k_i}\right](1 - B)^{k_o} T(X_t)$$

Here B is the familiar backshift operator of Reference [1]; i.e., $Bz_t = z_{t-1}$. The first group of factors in the above expression represent seasonal differences, the i^{th} factor denoting k_i seasonal differences of order o_i, where $o_i \geq 2$ to distinguish seasonal differences from regular differences. The $(1 - B)^{k_o}$ factor indicates k_o regular differences. Finally, $T(X_t)$ is a transformation that may either be of the form $T(X_t) = \log(X_t + \tau)$ or $T(X_t) = (X_t + \tau)^\alpha$, $0 \leq \alpha \leq 1$. The constant τ simply insures that $X_t + \tau \geq 0$, for all t.

The instruction keywords for regular differencing, seasonal differencing, and transformation are, respectively, RDIF, SDIF, and TRAN. The RDIF instruction has the format:

RDIF [nrdiff]

Nrdiff is the number of regular differences. The SDIF instruction is

SDIF ord, [nsdiff]

Ord is the order of the seasonal difference factor; that is, the integer quantity o_i in the expression $\left(1 - B^{o_i}\right)$. *Nsdiff* is the number of seasonal differences. It should be clear that no more than one RDIF instruction will be necessary; however, as many as four SDIF instructions are permitted in a program run. The TRAN instruction format is

TRAN [alpha] , [tau]

Tau is any quantity τ such that $X_t + \tau \geq 0$ for all t . *Alpha* is a value lying in $[0,1]$. If some value $\alpha > 0$ is chosen, then the input series will be transformed according to $T(X_t) = (X_t + \tau)^\alpha$. If *alpha* is zero, the transformation is $T(X_t) = \log(X_t + \tau)$.

With regard to the estimation and forecasting stages, these operations represent the initial part of model specification; specification is completed by choosing a suitable ARMA model for the x_t series. The RDIF instruction is not effective in the identification stage, but SDIF and TRAN instructions will allow the user to identify the series that results from operating on the input series.

Model Specification

The purpose of the identification stage is to give the user sufficient information about the input series to select an appropriate ARIMA model which describes the process. A set of instruction cards which specify the form of this model is required for estimation and forecasting. The discussion here will center on the procedure for specifying an ARMA model for the stationary series x_t obtained from the input series through some differencing and transformation scheme. If differencing and transformation are unnecessary, then x_t will denote the input series itself.

The TIMES program can accommodate a general ARMA model of the form:

$$\prod_{i=1}^{n_\phi} \phi_i(B)(x_t - \mu) = \theta_0 + \prod_{j=1}^{n_\theta} \theta_j(B)a_t$$

Here μ is the mean parameter and θ_0 is the trend parameter. Only one of these two parameters will appear in a particular model; if x_t is the result of differencing the input series, then no mean is allowed and the trend parameter is optional; otherwise, no trend is allowed, and a mean parameter is required. The terms $\phi_i(B)$ and $\theta_j(B)$ represent, respectively, the i^{th} autoregressive factor and the j^{th} moving average factor. They are polynomials in the backshift operator B :

$$\phi_i(B) = \left(1 - \phi_{a_i} B^{a_i} - \phi_{b_i} B^{b_i} - \ldots - \phi_{c_i} B^{c_i}\right)$$

$$\theta_j(B) = \left(1 - \theta_{q_j} B^{q_j} - \theta_{r_j} B^{r_j} - \ldots - \theta_{s_j} B^{s_j}\right)$$

Finally, a_t is a pure (white) noise series.

A mean or trend parameter is indicated by an instruction card bearing the keyword MEAN or TRND. Both cards have the same format:

$$\left.\begin{array}{l}\text{MEAN}\\ \text{TRND}\end{array}\right\} \text{[initv]}$$

For the estimation stage, *initv* will be the initial parameter value, giving one parameter space coordinate location where the nonlinear least squares minimization search is to begin. If the forecasting stage is performed in the same program run as estimation, the final parameter value resulting from the minimization process will be passed automatically to the forecasting subroutine. However, if forecasting is done in a separate run from estimation, *initv* must be the final parameter value determined from the estimation stage.

The keywords AR and MA are used for autoregressive and moving average parameters, respectively. A parameter of either type can be uniquely identified by giving the index of the factor to which it belongs and the power of B which it multiplies in that factor. The indexing scheme for the factors is quite arbitrary, the only restrictions being that each factor must have an integer index between 1 and 5, and no two factors of the same type (autoregressive or moving average) may have the same index. Thus up to 5 autoregressive factors and 5 moving average factors may appear in the same model. The integer power of B multiplied is the entry *ord* in the following format, and *ifac* is the factor index:

$$\left.\begin{matrix}\text{AR}\\ \text{MA}\end{matrix}\right\}\ \text{[initv],ord, ifac}$$

Initv is again the initial parameter value, and the remarks above concerning the same entry on MEAN and TRND instructions also apply here.

<u>A Specification Example</u>

For illustration, suppose the model is given formally as:

$$(1 - B^7)(1 - \phi B)\log(X_t) = \theta_0 + (1 - \Theta B^7)(1 - \theta B)a_t \ .$$

The TIMES instructions necessary to specify this model, with initial parameter values are

```
SDIF  7
TRAN  0
AR    .1,1,1
TRND  .2
MA    .5,1,1
MA    .8,7,2
```

Estimation

The EST instruction informs the control program that routines for parameter estimation are to be executed. The format is:

EST [nracor], [rplot], [nrchi], [niter], [eps1], [eps2]

After estimates of model parameters have been found, the residual series is calculated and automatically analyzed by the same routines utilized in the identification stage. Thus the entries *nracor, rplot,* and *nrchi* have the same significance with respect to the residual series as the entries *nacor, dplot,* and *nchi* described above. *Niter* gives the maximum number of iterations permitted in the nonlinear (Marquardt) minimization routine. *Eps1* is the relative error in the sum of squares value that is tolerable when terminating the minimization routine. Similarly, *eps2* is the relative error tolerable in the estimated value of each model parameter when terminating the minimization routine. Output from the estimation phase consists of a summary of the status of each iteration of the minimization routine, followed by the estimated parameter values. A cumulative periodogram and histogram of the residuals are also provided, as well as information necessary for the identification of the residuals.

Forecasting

The forecasting stage is controlled by a FCST instruction:

FCST [conf], nfcsts, [to$_1$], [to$_2$], [to$_3$], [to$_4$], [to$_5$]

Conf may be one of the integers 1, 2, 3, 4, or 5, which indicate that, respectively, 50%, 75%, 90%, 95%, or 99% confidence limits are to be provided with the forecast. *Nfcsts* is the number of forecasts to be made forward from each of the time origins $to_1, \ldots, to_5$. Though the entry for each time origin is optional, forecasts are done only for consecutive entries. For instance, the following card will result in forecasts only from time origins 120 and 140:

```
FCST  3, 15, 120, 140,,160
```

The UPDA instruction permits updated forecasting from the last origin entered on the FCST card:

```
UPDA  nupts, [iuplot]
```

This instruction may only be used once in a program run and must be accompanied by a FCST card. *Nupts* is the total number of data points to be used in updating; it should be noted that the sum of the last time origin on the FCST card and *nupts* may not exceed the number of data points, *npts*, in the data package. For the i^{th} update iteration, points $x_1, \ldots, x_{to_\ell+i}$ are assumed to be known, where to_ℓ is the last time origin. Forecasts are calculated for time indices $to_\ell + i + 1, to_\ell + i + 2, \ldots, to_\ell + nfcsts$. Forecast values and confidence limits are printed out for each update iteration; but plotting of the forecasts is optional, since it will usually be expensive and unnecessary to produce a new graph for each additional point used in updating. *Iuplot* is an integer between 1 and *nupts* that controls the number of forecast plots that are to be made. If *iuplot* is 4, for example, a plot is made after each group of four successive new data points has been incorporated into an updated forecast. A plot is always made after all *nupts* have been utilized.

Multiple Runs

The RERUN instruction card has no entry list and consists only of the keyword 'RERUN' in columns 1 through 5. The main purpose of this card is to allow the estimation and forecasting stages to be performed for a number of different time series models within the same computer job. The instruction reinitializes all control program switches and storage locations except those associated with the DATA and NAME instructions. Thus, following a RERUN instruction, a new set of model specification cards may be provided as well as EST and FCST instructions; and the time series routines will again be executed, utilizing the original data package and project name. If a NAME or DATA instruction is used following the RERUN instruction, the new name or data package will replace the original information.

III. SINGLE INPUT TRANSFER FUNCTION MODELLING

The objective of transfer function modelling is to investigate the relationship between two time series. It is assumed that the *dependent* series, denoted here by y_t , can be represented by a linear operation on the *independent* series, x_t , plus a noise component n_t , where y_t , x_t , and n_t are all stationary. In general, x_t and y_t will be generated by differencing and possibly transforming the actual data series X_t and Y_t . Hopefully, additional information available from the X_t series can be used to obtain better forecasts of Y_t than would be the case were Y_t treated alone. The justification and motivation for the techniques considered here are developed in some detail in Chapter 11 of Reference [1].

Transfer function modelling involves the familiar stages of identification, estimation, and forecasting. These transfer function stages are handled by the control program in a manner that parallels the single-series case. The procedures discussed below differ from those of Part II mainly in the complexity of model specification.

Data Input

The DATA and NAME instruction of Part II are to be used for entering information concerning the X_t series. YDATA and YNAME are the parallel instructions associated with the Y_t series; these instructions have the same format and are used in the same manner as DATA and NAME. Should the number of data points in X_t differ from the number in Y_t , the minimum of the two series lengths will be used for analysis, with the remaining points in the longer series ignored.

The data points for both series are required in each stage of analysis. The NAME and YNAME instructions are optional.

Differencing and Transformation

The scheme of regular and seasonal differencing and data transformation necessary to generate the stationary series x_t and y_t from the original series X_t and Y_t should be determined prior to transfer function identification. Differencing and transformation for X_t can be determined when this series is analyzed by the procedures of Part II to obtain the prewhitening model, which is discussed below. To select the proper differencing and transformation for Y_t , it may be desirable to use Y_t as input to the single series identification stage of Part II.

The instructions RDIF, SDIF, and TRAN introduced earlier operate on the X_t series in transfer function modelling. The corresponding instructions for the Y_t series have the keywords YRDIF, YSDIF, and YTRAN, and the formats of these cards are the same, respectively, as those of RDIF, SDIF, and TRAN. Each of the instructions RDIF, YRDIF, TRAN, and YTRAN may be used only once. However, up to four SDIF instructions and four YSDIF instructions are permitted in a run.

Prewhitening Model Specification

Transfer function modelling involves three distinct model specifications: the prewhitening model, the noise model, and the transfer function. The results of the transfer function identification stage will suggest the forms of the noise model and the transfer function; parameter values for these two models will then be determined simultaneously in the estimation stage. The prewhitening model, on the other hand, must be selected prior to the identification stage by considering X_t alone in the analysis described in Part II. The residuals from applying the prewhitening transformation to x_t should approximate a pure noise series, and this is in fact the objective of single series analysis. The model specification discussed in Part II, utilizing the keywords AR, MA, MEAN, and TRND, applies without modification to the prewhitening model. The *initv* entry for each parameter card should be the final estimated value obtained for that parameter.

The same prewhitening model instructions should be provided in each stage of transfer function analysis. If no prewhitening model instructions are used, the prewhitening transformation will be the identity.

Identification

The stationary series x_t and y_t will result from differencing and transforming the input series X_t and Y_t ; if differencing is unnecessary, x_t and y_t will be the original or transformed series translated so that the means are zero. (The routines perform this translation automatically.) As discussed in Reference [1], the transfer function model in its basic form can be written as

$$y_t = \gamma + v(B)x_t + n_t \ ,$$

where n_t is a stationary noise series with zero mean, γ is a possible trend parameter, and $v(B)$ is an operator of the form

$$b(B) = \left(v_0 + v_1 B + v_2 B^2 + \ldots\right) \ .$$

The purpose of the identification stage is to provide rough estimates of the v weights and the noise series. If the estimated noise series has a nonzero mean, then this mean can be taken as an initial estimate of γ .

When the prewhitening transformation is applied to x_t , an approximate pure noise series α_t will result. The key to estimating the v weights is to apply the same transformation to y_t to obtain a residual series β_t ; it is shown in Reference [1] that if $r_{\alpha\beta}(k)$ is the sample crosscorrelation function between α_t and β_t , and s_β and s_α are the sample variances, then

$$\hat{v}_k = \frac{r_{\alpha\beta}(k)s_\beta}{s_\alpha} \qquad k = 0,1,2, \ldots$$

An estimate of the noise series n_t may then be calculated according to the relation

$$\hat{n}_t + \hat{\gamma} = y_t - \left(v_0 + v_1B + v_2B^2 + \ldots\right)x_t .$$

The TIDEN instruction card initiates the above manipulations and has the format:

TIDEN [nccor], [dplot], [nvwts], [nnacor], [ndplot], [nnchi]

Nccor is the number of lags for which the sample crosscorrelation function $r_{\alpha\beta}(\cdot)$ is to be calculated. If *dplot* is 1, a combined plot of the original series X_t and Y_t is provided. The entry *nvwts* specifies the number of v-weight estimates to be utilized in calculating the noise series. The estimated noise series is analyzed by the single series identification routine of Part II. Thus *nnacor*, *ndplot*, and *nnchi* correspond, respectively, to the entries *nacor*, *dplot*, and *nchi* on the IDEN instruction card.

Transfer Function Specification

The v weights are used only indirectly in formulating the transfer function model; in the simplest case, these weights are taken to be functions of a small number of parameters associated with an *input lag factor*, $\omega(B)$, and an *output lag factor*, $\delta(B)$. These factors are of the form

$$\omega(B) = \left(\omega_0 - \omega_1B - \ldots - \omega_pB^p\right)$$

$$\delta(B) = \left(1 - \delta_1B - \ldots - \delta_qB^q\right) ,$$

and $v(B)$ is assumed to have the representation

$$v(B) = \delta^{-1}(B)\omega(B)B^{b}$$

Here b will be greater than zero if $v_0 = v_1 = \ldots = v_{b-1} = 0$. In general, $\omega(B)$ and $\delta(B)$ are to be chosen with as few parameters as will allow the coefficients of B^k in $\delta^{-1}(B)\omega(B)B^b$ to approximate the estimated sequence $\hat{v}_k$ determined in the identification stage. The process of selecting input and output lag factors may require considerable judgment and care on the part of the analyst. This process will also yield initial estimates of parameters ω_i , $i = 1, \ldots, p$ and δ_j, $j = 1, \ldots, q$.

The estimation and forecasting routines, in fact, can accommodate multiple lag factors, since seasonal relationships between y_t and x_t may require this additional flexibility. Thus the general model is as follows:

$$y_t = \gamma + \prod_{j=1}^{n_\delta} \delta_j^{-1}(B) \prod_{i=1}^{n_\omega} \omega_i(B)x_{t-b} + n_t \ ,$$

where

$$\omega_i(B) = \left(1 - \omega_{a_i}B^{a_i} - \omega_{b_i}B^{b_i} - \ldots - \omega_{c_i}B^{c_i}\right) , \qquad i \neq 1$$

$$\omega_1(B) = \left(\omega_{0_1} - \omega_{q_1}B^{q_1} - \omega_{r_1}B^{r_1} - \ldots - \omega_{s_1}B^{s_1}\right)$$

$$\delta_j(B) = \left(1 - \delta_{u_j}B^{u_j} - \delta_{v_j}B^{v_j} - \ldots - \delta_{w_j}B^{w_j}\right)$$

The number of factors of either type is limited to five. Each transfer function parameter can be uniquely identified by the index of the lag factor to which it belongs and the power of B which it multiplies. It should be noted that the only parameter that does not multiply a positive power of B

is ω_{0_1} ; that is, this parameter must be in the first input lag factor. With this exception, the indexing of the lag factors is unimportant.

The instruction cards for the input lag and output lag parameters have the respective keywords ILAG and OLAG:

ILAG
OLAG [initv], ord, ifac

Initv is the initial value of the parameter for estimation. As before, if forecasting is done in a separate run from estimation, *initv* must be the value of the parameter from estimation. *Ord* is the power of B that the parameter multiplies, and *ifac* is a number between 1 and 5 giving the index of the factor to which the parameter belongs.

Transfer function model specification is completed by optionally specifying the trend parameter γ and the number of lags *b* by which perturbations in y_t lag those in x_t . These two instructions are as follows:

TTRND [initv]
B b

Up to ten B instructions are permitted in an estimation run. Parameter estimation will be repeated for each *b* value.

Noise Model Specification

The noise series n_t is assumed to be stationary and have zero mean, but in most cases it will not be a pure noise series. An analysis of the sample autocorrelation and partial autocorrelation functions of the estimated series $\hat{n}_t$ calculated in the identification stage will suggest an appropriate model for n_t in terms of autoregressive and moving average factors. The general form of the model is

$$\prod_{i=1}^{n_\phi} \phi_i(B)n_t = \prod_{j=1}^{n_\phi} \theta_j(B)a_t ,$$

where a_t is a pure noise series and

$$\phi_i(B) = \left(1 - \phi_{a_i} B^{a_i} - \phi_{b_i} B^{b_i} - \ldots - \phi_{c_i} B^{c_i}\right)$$

$$\theta_j(B) = \left(1 - \theta_{q_j} B^{q_j} - \theta_{r_j} B^{r_j} - \ldots - \theta_{s_j} B^{s_j}\right) .$$

Thus each parameter in the noise model is either an autoregressive or a moving average parameter and is uniquely identified by the index of the factor to which it belongs and the power of B which it multiplies. The keywords are NAR and NMA for autoregressive and moving average parameters, respectively, and the instruction card formats are

```
NAR  [initv], ord, ifac
NMA
```

Initv, ord, and *ifac* are defined in the usual manner, and up to five factors of either type are permitted.

A Specification Example

An example is appropriate to illustrate the use of model specification instructions. The structure of the two series considered in Section 11.5.3 of Reference [1] may be expressed as

$$y_t = .035 + (1 - .72B)^{-1}(4.82)x_{t-3} + (1 - .54B)a_t$$

$$x_t = (1 - .32B)\alpha_t ,$$

where y_t and x_t are first differences of the original series and a_t and α_t are pure noise series. The complete set of instruction cards for all three models, plus differencing instructions, is as follows:

```
RDIF
YRDIF
MA     .32,1,1
OLAG   .72,1,1
ILAG   4.82,0,1
B      3
NMA    .54,1,1
```

Estimation

The estimation stage is very similar to the single series estimation of Part II. The prewhitening, transfer function, and noise models should be specified, with initial estimates for transfer function and noise parameters. The instruction format for estimation is

```
TEST  [nrccor], [nracor], [rdplot], [nrchi], [niter], [eps1], [eps2]
```

The first five entries are associated with the analysis of the residuals, $\hat{a}_t$, from the estimated noise series $\hat{n}_t$, which is determined after final parameter estimates have been obtained. *Nrccor* is the desired number of lags for which the sample crosscorrelation function between a_t and the prewhitened x_t series, α_t , is calculated. As explained in Section 11.3.3 of Reference [1], this crosscorrelation is a measure of the discrepancy between the "correct" v weights for the transfer function model and the v weights that are calculated from the parameter estimates. The transfer function estimation routine utilizes the single series identification routine to provide information for analysis of the residuals; so again, *nracor, rdplot,* and *nrchi* correspond to the entries for the IDEN instruction. Finally, *niter, eps1,* and *eps2* are control values for terminating the least-squares minimization process; they are considered in Part II in conjunction with the EST instruction.

Forecasting

For the forecasting stage, all three models should be specified, and if this stage is run separately from estimation, the *initv* entry for each transfer function and noise parameter must be the final estimate for that parameter. TFCST is the forecasting instruction:

TFCST [conf], [nfcsts], [to_1], [to_2], [to_3], [to_4], [to_5]

The entries for this instruction correspond exaccly to the entries on the FCST instruction card. The single-series forecasting routine automatically provides forecasts of the X_t series, utilizing the prewhitening model. The X_t forecasts are then taken as input to the transfer function model to eventually yield forecasts of Y_t .

The single-series option of updated forecasting is not available in the transfer function forecasting stage.

Multiple Runs

The RERUN instruction may also be used in transfer function modelling. The instruction reinitializes the control program, but leaves data and names for both series unchanged.

IV. SPECTRAL IDENTIFICATION

Spectral methods of time series analysis differ significantly from the Box-Jenkins approach, but these methods are often quite useful, especially for time series arising from natural phenomena, such as seismic, atmospheric, or biological processes. Detailed discussions of spectral techniques may be found in Brillinger [2] and Watts and Jenkins [3]. The TIMES program has options for both the spectral analys_s of a single time series and the cross-spectral analysis of a pair of series.

Let $c_{xx}(\cdot)$ be the autocovariance function of a time series x_t . The *power spectrum* of this series is the Fourier transform of the autocovariance function:

$$f_{xx}(w) = \frac{1}{2\pi} \sum_{k=-\infty}^{\infty} c_{xx}(k)e^{-i2\pi kw} \text{ , } 0 \leq w \leq .5 \text{ , } i = \sqrt{-1}$$

Over any finite time interval, the process x_t may be approximated by a process $\hat{x}_t$ consisting of a sum of sinusoidal components having random amplitudes and phases. These components are uncorrelated and have frequencies in the range -.5 to .5 cycles per time unit. (A time unit is the duration between successive points of the x_t series.) The approximation can be made arbitrarily close in the sense that $E(x_t - \hat{x}_t)^2$ can be made small. As $E(x_t - \hat{x}_t)^2 \rightarrow 0$, these components may be thought of as a decomposition of the x_t process, and the power spectrum represents the squared component magnitudes as a function of the frequency w .

In Part III the dependent series y_t was assumed to be related to x_t by

$$y_t = \gamma + v(B)x_t + n_t \text{ ,}$$

where γ is a possible trend parameter, $v(B)$ is the generating function of the v-weights, and n_t is a stationary noise series, uncorrelated with x_t. Let $c_{yx}(\cdot)$ be the cross covariance function between y_t and x_t; that is, $c_{yx}(k) = \mathrm{Cov}[y_{t+k}, x_t]$. The *cross spectrum* is the Fourier transform of the cross covariance:

$$f_{yx}(w) = \frac{1}{2\pi} \sum_{k=-\infty}^{\infty} c_{yx}(k) e^{-i2\pi kw} \quad -.5 \leq w \leq .5 .$$

It is not difficult to show that

$$v(e^{-i2\pi w}) = \sum_{k=-\infty}^{\infty} v_k e^{-i2\pi kw} = \frac{f_{yx}(w)}{f_{xx}(w)} .$$

Now $v(e^{-i2\pi w})$ may be written as the product of a real function $G(w)$, called the *gain function*, and $e^{i2\pi\phi(w)}$, where $\phi(w)$ is a (real) *phase function:*

$$v(e^{-i2\pi w}) = G(w) e^{i2\pi\phi(w)}$$

The sinusoidal components associated with $v(B)x_t$ can essentially be obtained by multiplying the components of x_t by $v(e^{-i2\pi w})$. This multiplication alters the magnitude of the component with frequency w through the factor $G(w)$; the phase is shifted by the factor $e^{i2\pi\phi(w)}$. The impulse response function may be evaluated from $v(e^{-i2\pi w})$ through an inverse Fourier transform:

$$v_k = \int_{-1/2}^{1/2} v(e^{-i2\pi w}) e^{i2\pi wk} dw$$

Another important function in cross-spectral analysis is the *coherence*, defined by

$$\kappa_{yx}^2(w) = \frac{|f_{yx}(w)|^2}{f_{xx}(w)f_{yy}(w)} .$$

Roughly, this quantity behaves like a squared correlation coefficient for each frequency w, measuring the correlation between the sinusoidal component of y_t and that of x_t at the frequency w. Furthermore, it can be shown that the spectrum of the noise series may be evaluated as

$$f_{nn}(w) = f_{yy}(w)\left[1 - \kappa_{yx}^2(w)\right] .$$

Thus $\kappa_{yx}^2(w)$ is in some sense a measure of the portion of information in the y_t series that is directly attributable to the x_t series. Finally, the noise autocovariance function may be obtained by taking the Fourier inverse of $f_{nn}(\cdot)$:

$$c_{nn}(k) = 2 \int_0^{1/2} f_{nn}(w)\cos(2wk)dw .$$

For a single series x_t, the TIMES spectral analysis routine simply estimates the power spectrum $f_{xx}(\cdot)$. For a pair of series, x_t and y_t, the routine first estimates $f_{xx}(\cdot)$, $f_{yy}(\cdot)$, and $f_{yx}(\cdot)$. These estimates are used in turn to estimate the gain, phase, coherence, and noise spectrum, in a manner suggested by the above definitions of these functions in terms of $f_{xx}(\cdot)$, $f_{yy}(\cdot)$, and $f_{yx}(\cdot)$. From the gain and phase estimates the impulse response weights are determined by Fourier inversion, and the noise spectrum estimate is inverted to yield the noise autocorrelation function.

Estimates of $f_{xx}(\cdot)$, $f_{yy}(\cdot)$, and $f_{yx}(\cdot)$ are based on the *discrete Fourier transform,* defined for an arbitrary (real or complex) time series z_t, $t = 0, \ldots, T - 1$, by

$$d_z^T(w) = \sum_{t=0}^{T-1} z_t e^{-i2\pi tw} \, , \; -.5 \le w \le .5 \, .$$

The Fast Fourier Transform algorithm is used to compute the sample discrete transform, so a particular realization of $d_z^T(\cdot)$ is calculated only for 2^N equally spaced frequencies $w = j/2^N - .5$, $j = 1,2, \ldots, 2^N$, where N is chosen such that $2^N \geq T$. For the series x_t and y_t , both of length T , *periodograms* $I_{xx}(\cdot)$ and $I_{yy}(\cdot)$ are defined by

$$I_{xx}(w) = \frac{1}{2\pi T} d_x^T(w)\overline{d_x^T(w)} \, ,$$

$$I_{yy}(w) = \frac{1}{2\pi T} d_y^T(w)\overline{d_y^T(w)} \, , \quad 0 \le w \le .5 \, ,$$

where the bar denotes conjugation. The *cross periodogram* is

$$I_{yx}(w) = \frac{1}{2\pi T} d_y^T(w)\overline{d_x^T(w)} \, , \quad -.5 \le w \le .5 \, .$$

Now $I_{xx}(w)$, $I_{yy}(w)$, and $I_{yx}(w)$ are random variables that have respective expected values $f_{xx}(w)$, $f_{yy}(w)$, and $f_{yx}(w)$, as shown in Reference [2]; however, these periodogram ordinates have variances that do not decrease as T increases. More stable estimators $f_{xx}^T(w)$, $f_{yy}^T(w)$, and $f_{yx}^T(w)$ are therefore obtained by smoothing the periodograms:

$$f_{xx}^T(w) = \frac{1}{(2m+1)} \sum_{k=-m}^{m} I_{xx}(w + k/2^N) \, ,$$

$$f_{yy}^T(w) = \frac{1}{(2m+1)} \sum_{k=-m}^{m} I_{yy}(w + k/2^N) \, ,$$

$$f^T_{yx}(w) = \frac{1}{(2m + 1)} \sum_{k=-m}^{m} I_{yx}(w + k/2^N) .$$

The quantity $(2m + 1)/2^N$ is the smoothing *bandwidth*, and it must be chosen large enough by the analyst so that spectral estimates are stable, but not so large as to obscure details of the spectra. It is good practice to perform the analysis several times using different bandwidths, most commonly on the range .01 to .1.

Reference [2] discusses the asymptotic distributions of the estimators $f^T_{xx}(\cdot)$, $f^T_{yy}(\cdot)$ and $f^T_{yx}(\cdot)$ and then derives various distributional properties for the coherence, phase, gain, and residual spectrum estimators. Approximate confidence limits for the estimates can therefore be calculated, and these approximations should be quite valid for data samples consisting of more than about 250 points. Estimates of power spectra and gain function are more appropriately given in terms of logarithms, since this simplifies computation of confidence limits. The TIMES spectral routine outputs $\log_{10}$ spectra and gain.

The spectral analysis subroutine is called if an SIDEN instruction is encountered in the control program input stream. The instruction format is

SIDEN [opt], [bw_1], [bw_2], [bw_3], [bw_4], [bw_5]

The entry *opt* determines the number of frequencies for which spectra, gain, phase, and coherence will be estimated. This entry does not determine the frequency range, which is always [0,.5] ; rather, it determines how finely this interval is to be divided for discrete output. If *opt* is 1, calculations are done for 101 frequencies. *Opt* may be 2 only if the data series passed to the spectral routine exceed 512 points in length, and in this case calculations are done for 251 frequencies. The entries bw_i, $i = 1, \ldots, 5$, are smoothing bandwidths. A complete spectral analysis will be done for each successive bandwidth entry.

Use of the SIDEN instruction will result in estimation of one power spectrum if the input stream contains only a single data package. If data packages for both the X_t and Y_t series are present, cross-spectral analysis will be done. The RDIF, SDIF, TRAN, YRDIF, YSDIF, and YTRAN instructions will apply differencing and transformation operations to the series prior to spectral analysis. In addition, a prewhitening model may be specified in the same manner as discussed in Part III; the prewhitening technique should, in fact, always be used when cross-spectral estimates are desired. Single-series spectral analysis gives the sample power spectrum of the residuals from applying the prewhitening model to the X_t series, so in single-series modelling the power spectrum is an additional tool to aid the analyst in identification of the residuals. If an SIDEN and an FST instruction are used in the same run, spectral analysis will follow the estimation stage in the processing sequence; thus, final parameter estimates, rather than the initial values on AR and MA instructions, will be used for the prewhitening transformation.

APPENDIX I

Instruction Entry Bounds and Defaults

The quantities NOBMX, NACMX, and NFMX are, respectively, the maximum number of observations, autocorrelations, and forecasts that are permitted in a particular implementation of the TIMES program. These quantities depend on the size of storage arrays in the control program; sizes may be altered only by changing several FORTRAN instructions in the program.

The abbreviation "req." indicates that the associated entry is required and no default value can be assigned. The abbreviation "succ." indicates that forecasting or spectral analysis will be done only for successive time origin entries or bandwidth entries. The "Multiple Usage" column contains a "yes" if two or more instructions having the associated keyword are permitted in a program run.

It should be noted that a run may not be successful even if all entries satisfy the bound constraints below: The control program offers only a preliminary check on the validity of some entries; final checks are done when the time series routines are executed.

Keyword	Multiple Usage	Entries	Lower Bound	Default	Upper Bound
DATA	no	npts	1	(req.)	NOBMX
NAME	no	(none)			
IDEN	no	nacor	1	npts/5	NACMX
		dplot	0	0	1
		nchi	1	nacor	NACMX
RDIF	no	nrdiff	1	1	100
SDIF	yes	ord	2	(req.)	10^{30}
		nsdiff	1	1	100
TRAN	no	alpha	0	0	1
		tau	-10^{30}	0	10^{30}

Keyword	Multiple Usage	Entries	Lower Bound	Default	Upper Bound
MEAN	no	initv	-10^{30}	.1	10^{30}
TRND	no	initv	-10^{30}	.1	10^{30}
AR	yes	initv	-1	.1	1
		ord	1	(req.)	10^{30}
		ifac	1	(req.)	5
MA	yes	initv	-1	.1	1
		ord	1	(req.)	100
		ifac	1	(req.)	5
EST	no	nracor	1	npts/5	NACMX
		rplot	0	0	1
		nrchi	1	nracor	NACMX
		niter	0	15	999
		eps1	10^{-8}	10^{-3}	.1
		eps2	10^{-8}	10^{-3}	.1
FCST	no	conf	1	3	5
		nfcsts	1	(req.)	NFMX
		to_i $(1 \leq i \leq 5)$	1	(succ.)	npts
UPDA	no	nupts	1	(req.)	NFMX
		iuplot	1	nupts	nupts
RERUN		(none)			
YDATA	no	npts	1	(req.)	NOBMX
YNAME	no	(none)			
YRDIF	no	nrdiff	1	1	100
YSDIF	yes	ord	2	(req.)	10^{30}
		nsdiff	1	1	100
YTRAN	no	alpha	0	0	1
		tau	-10^{30}	0	10^{30}

Keyword	Multiple Usage	Entries	Lower Bound	Default	Upper Bound
TIDEN	no	nccor	1	npts/5	NACMX
		dplot	0	0	1
		nvwts	1	nccor	NACMX
		nnacor	1	npts/5	NACMX
		ndplot	0	0	1
		nnchi	1	nnacor	NACMX
ILAG	yes	initv	-10^{30}	.1	10^{30}
		ord	0	(req.)	10^{30}
		ifac	1	(req.)	5
OLAG	yes	initv	-1	.1	1
		ord	1	(req.)	10^{30}
		ifac	1	(req.)	5
TTRND	no	initv	-10^{30}	.1	10^{30}
B	yes	initv	-10^{30}	(req.)	10^{30}
NAR	yes	initv	-1	.1	1
		ord	1	(req.)	10^{30}
		ifac	1	(req.)	5
NMA	yes	initv	-1	.1	1
		ord	1	(req.)	10^{30}
		ifac	1	(req.)	5
TEST	no	nrccor	1	npts/5	NACMX
		nracor	1	npts/5	NACMX
		rdplot	0	0	1
		nrchi	1	nracor	NACMX
		niter	0	15	999
		eps1	10^{-8}	10^{-3}	.1
		eps2	10^{-8}	10^{-3}	.1

Keyword	Multiple Usage	Entries	Lower Bound	Default	Upper Bound
TFCST	no	conf	1	3	5
		nfcsts	1	(req.)	NFMX
		to_i $(1 \leq i \leq 5)$	1	(succ.)	npts
SIDEN		opt	1	1	2
		bw_i $(1 \leq i \leq 5)$	0	(succ.)	.25

APPENDIX II

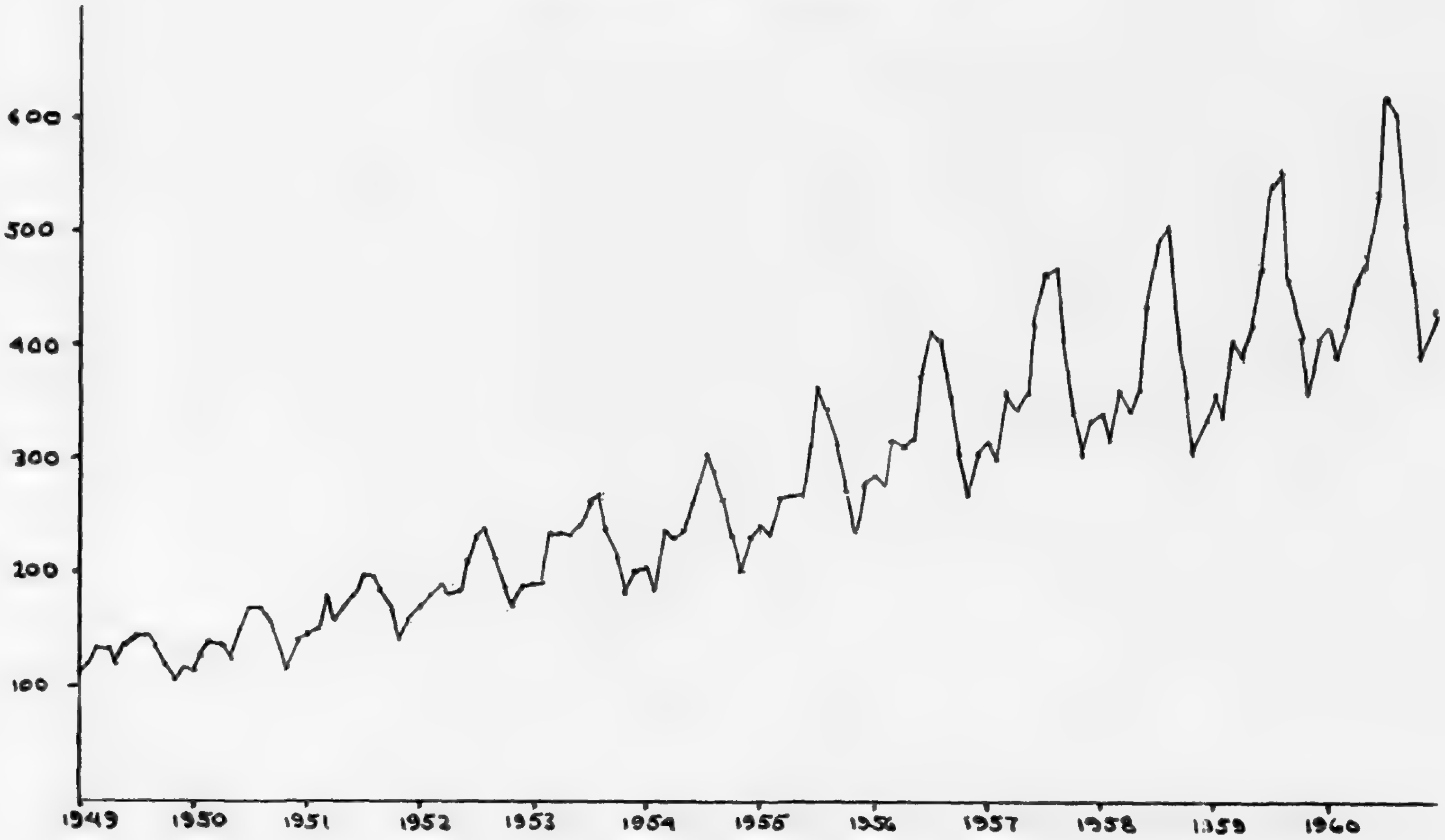
An Example of Modelling and Forecasting with a Single Time Series

Data for the following example is given in Reference [1], p. 531. The model for this particular series is considered in some detail in Chapter 9 of Reference [1].

INTERNATIONAL AIRLINE PASSENGERS

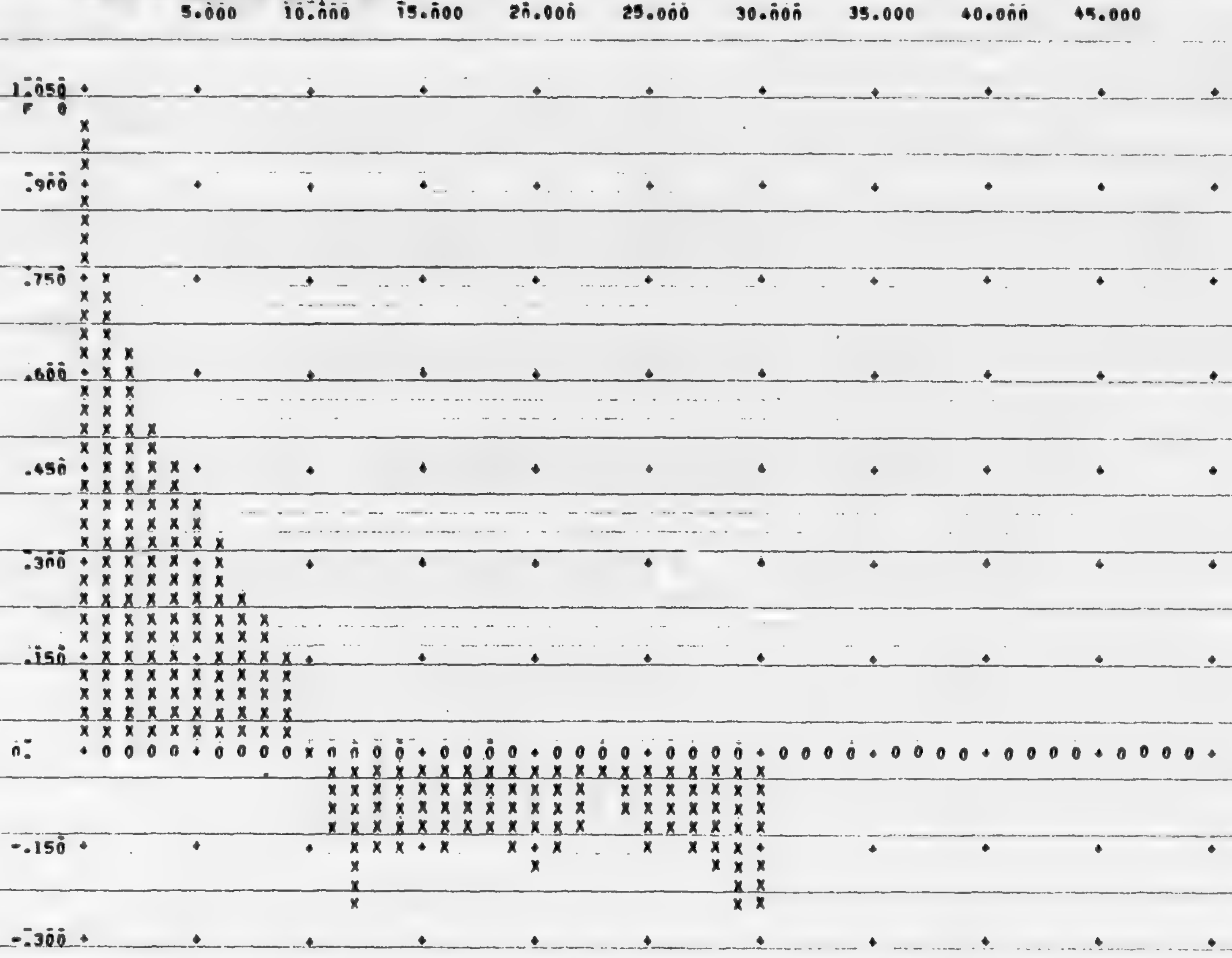
MONTHLY TOTALS JAN. 1949 - DEC 1960

(THOUSANDS OF PASSENGERS)

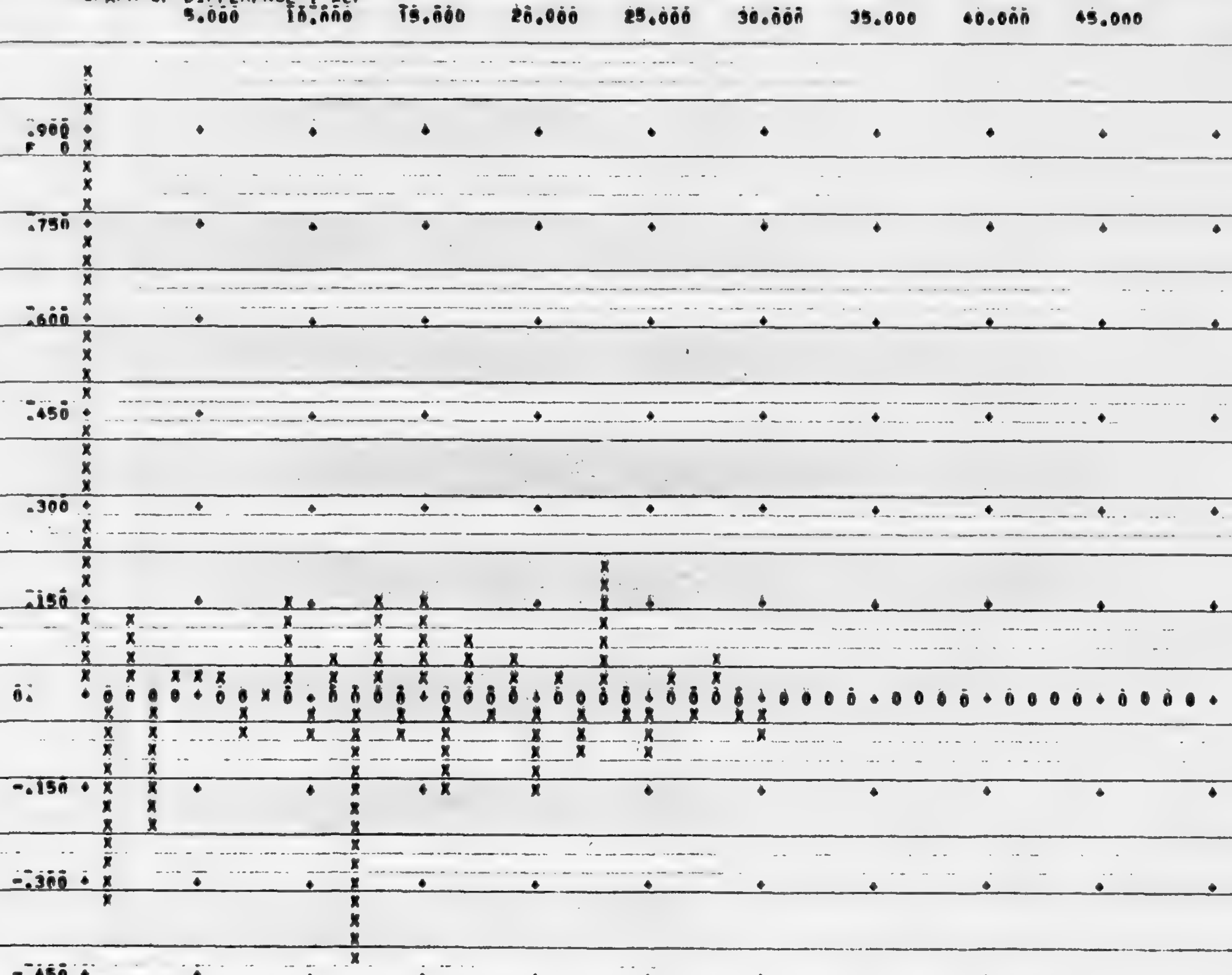


```
PROGRAM CONTROL STATEMENTS

          NAME
          DATA   132
          SDIF   12
          TRAN
          IDEN   30..30
```

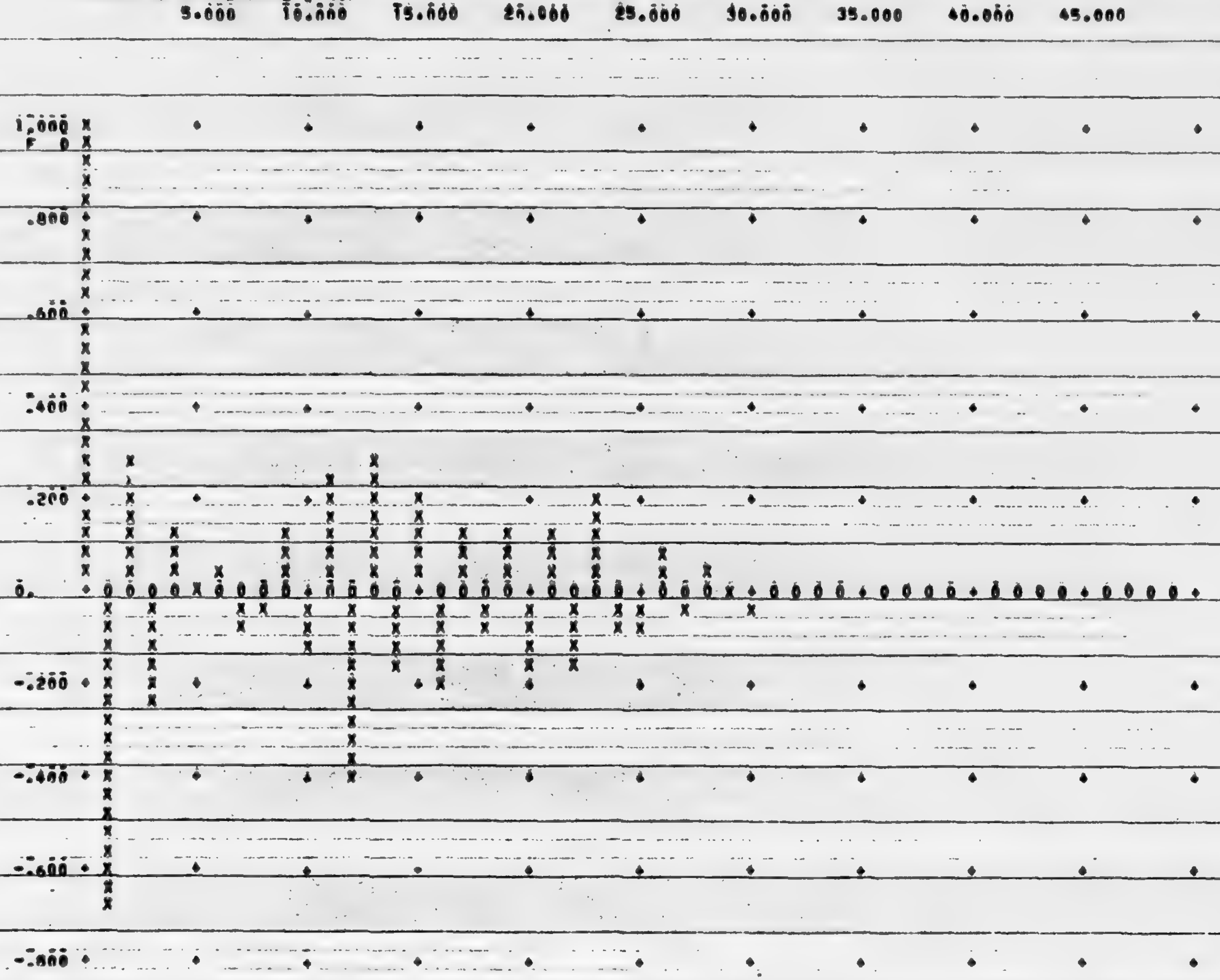
INTERNATIONAL AIRLINE PASSENGERS - MONTHLY TOTALS - JAN 1949 - DEC 1960
GRAPH OF OBSERVED SERIES ACF



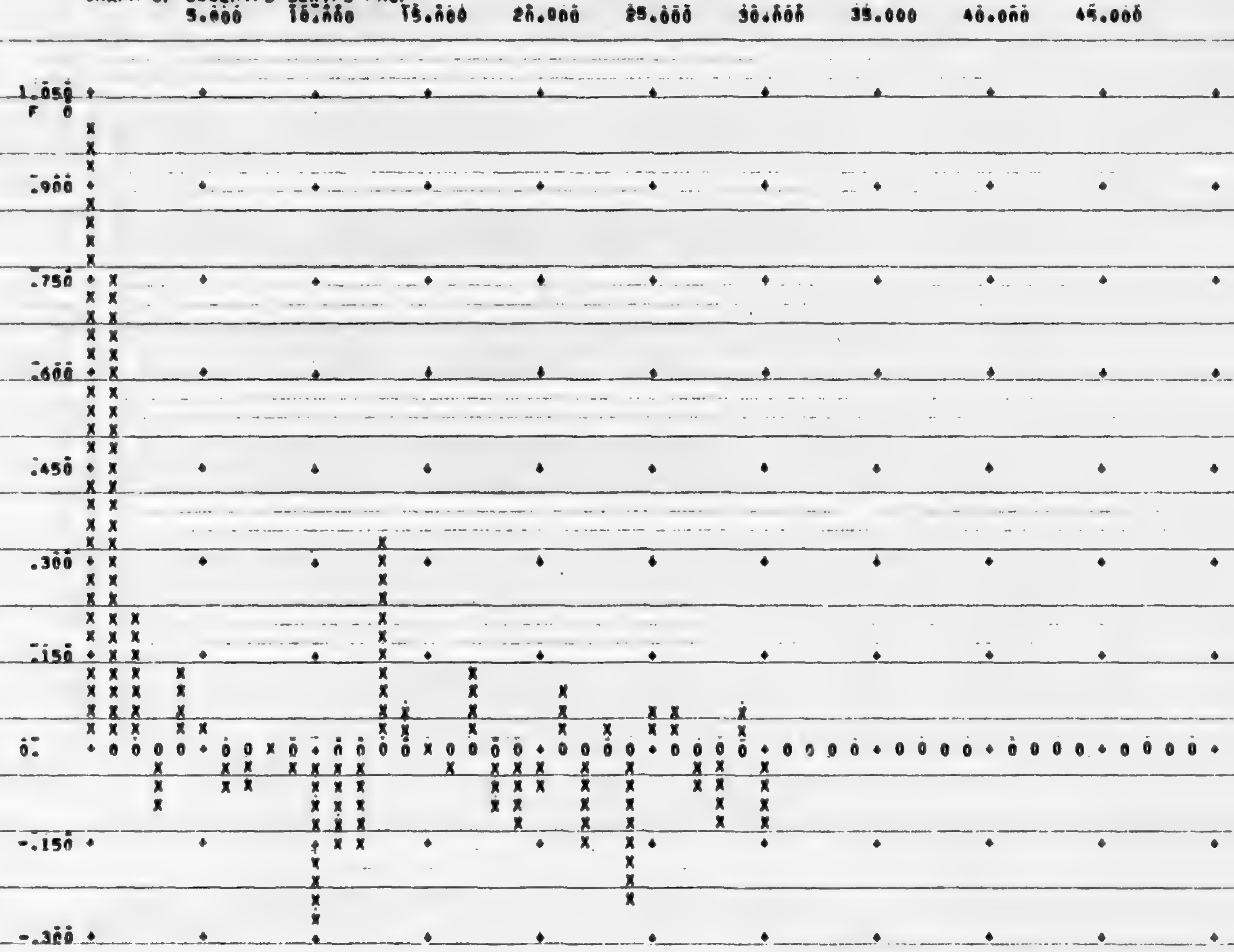
INTERNATIONAL AIRLINE PASSENGERS - MONTHLY TOTALS - JAN 1949 - DEC 1960
GRAPH OF DIFFERENCE 1 ACF



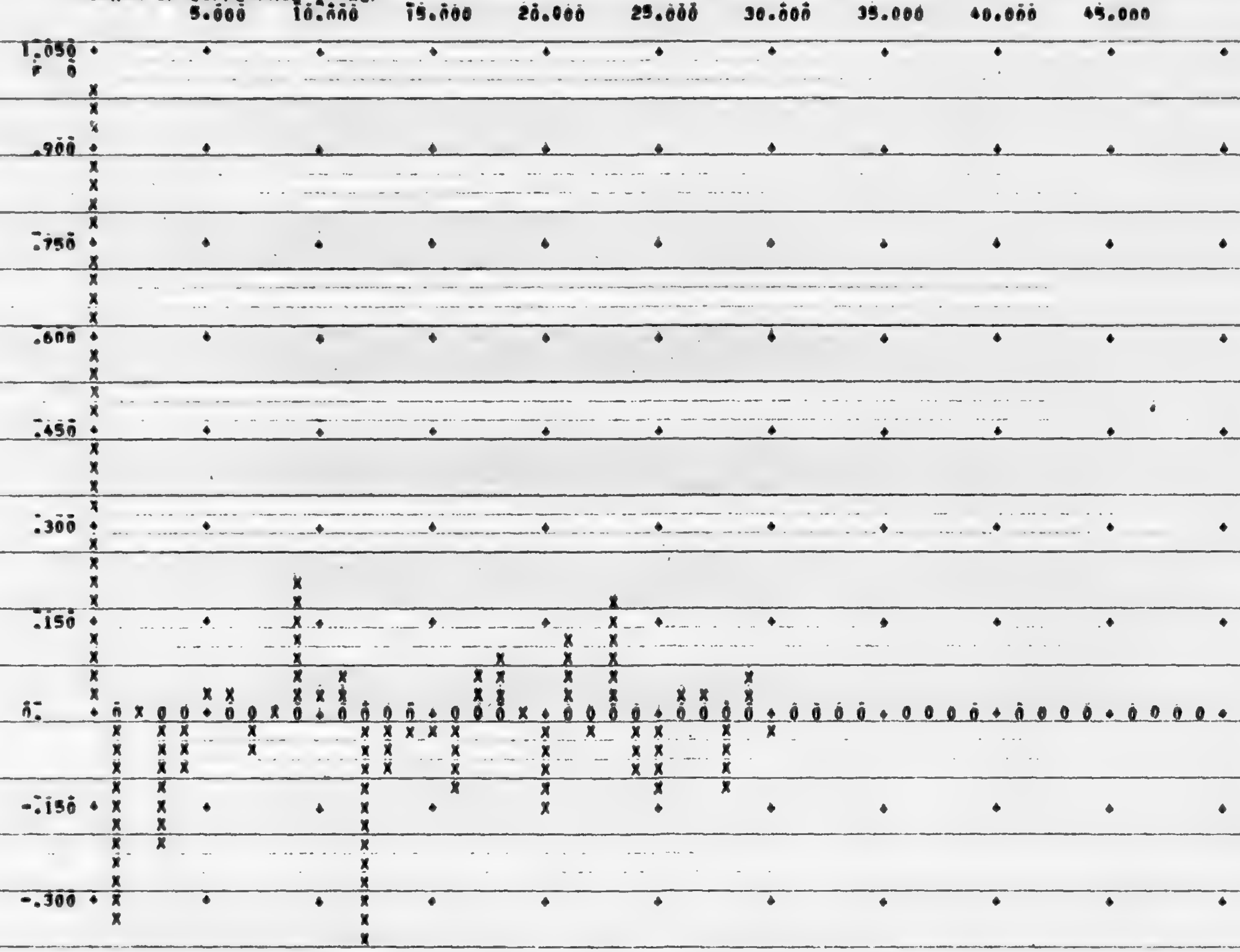
INTERNATIONAL AIRLINE PASSENGERS - MONTHLY TOTALS - JAN 1949 - DEC 1960
GRAPH OF DIFFERENCE 2 ACF



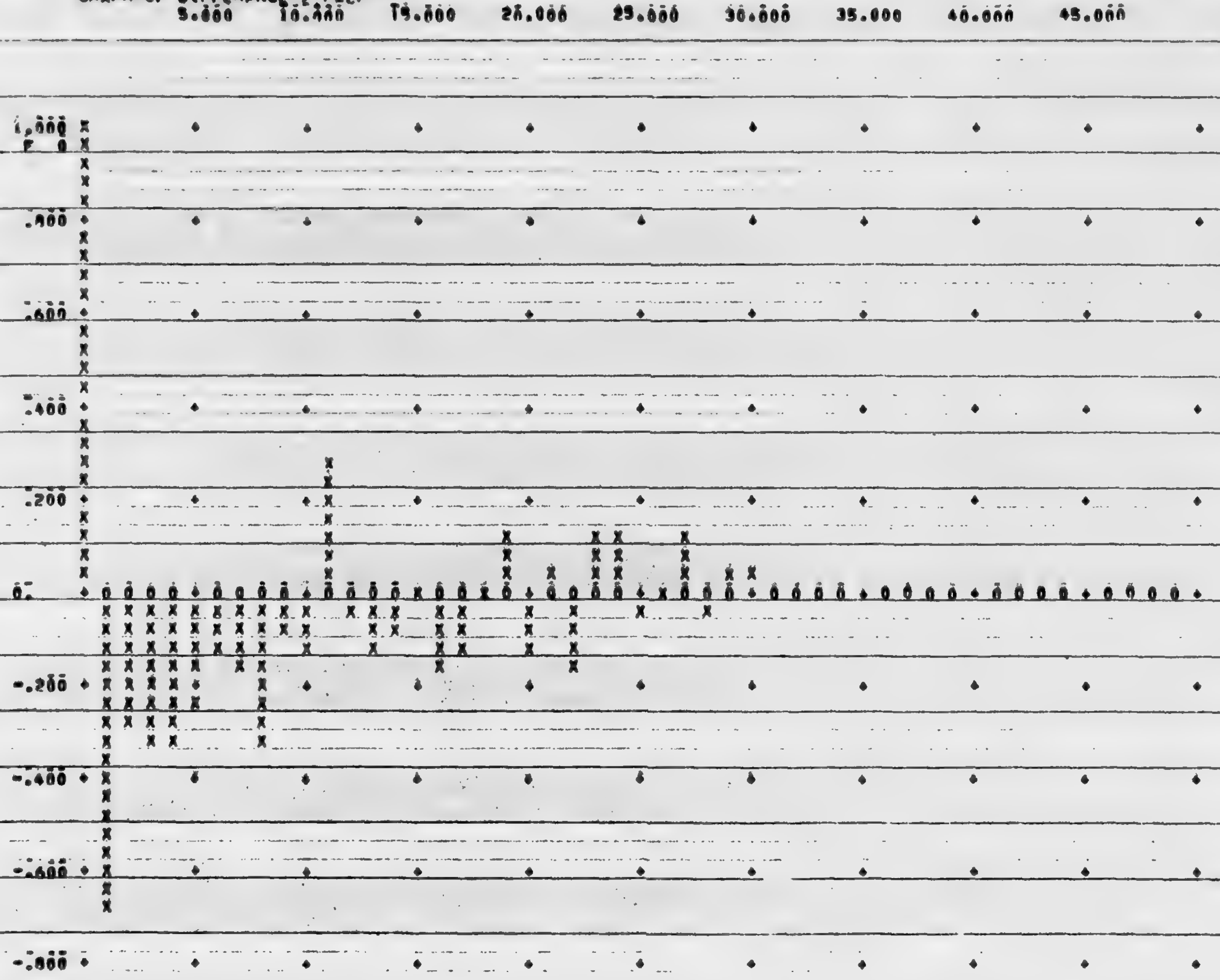
INTERNATIONAL AIRLINE PASSENGERS - MONTHLY TOTALS - JAN 1949 - DEC 1960
GRAPH OF OBSERVED SERIES PACF



INTERNATIONAL AIRLINE PASSENGERS - MONTHLY TOTALS - JAN 1949 - DEC 1960
GRAPH OF DIFFERENCE 1 PACF



INTERNATIONAL AIRLINE PASSENGERS - MONTHLY TOTALS - JAN 1949 - DEC 1960
GRAPH OF DIFFERENCE 2 PACF



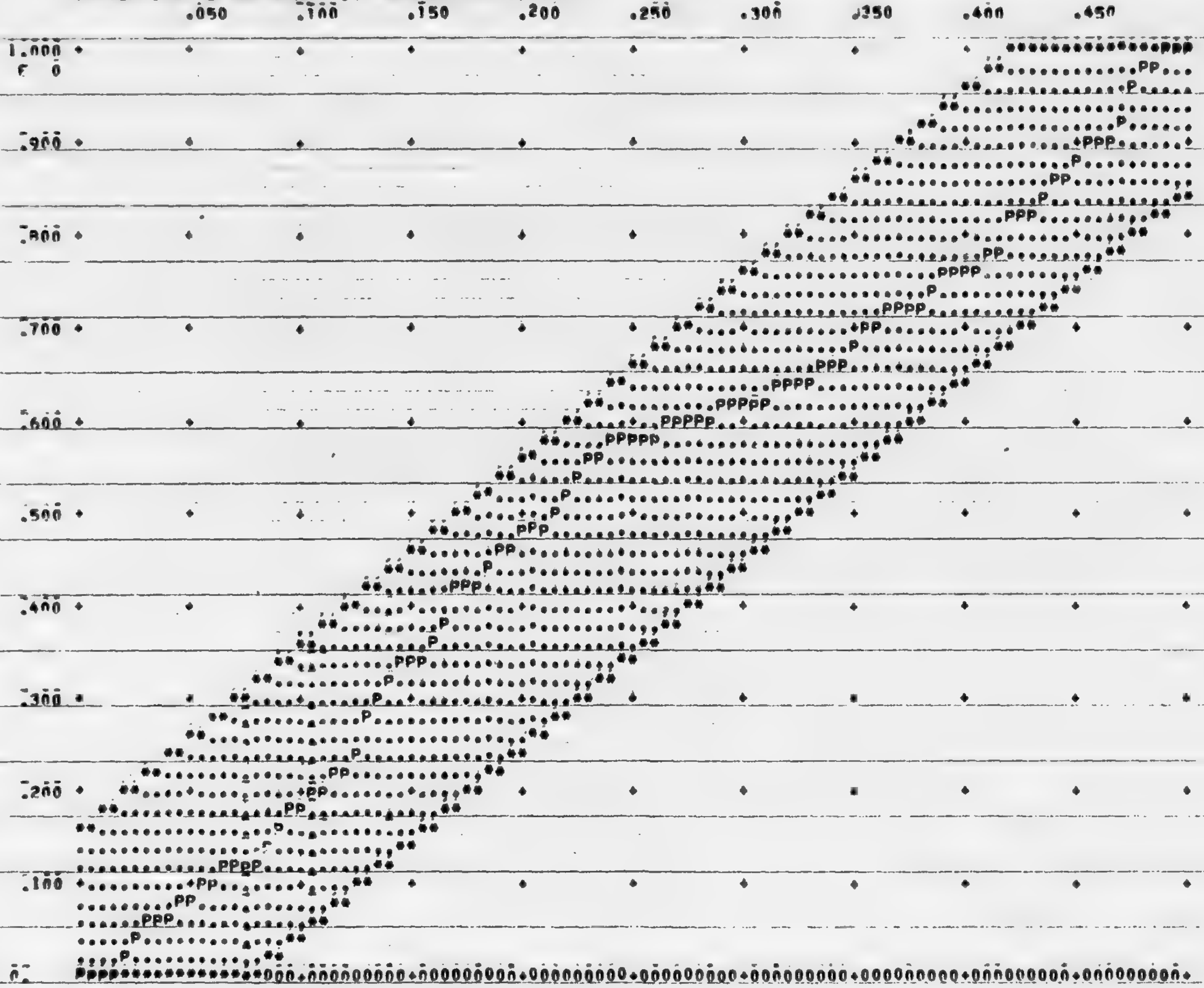
PROGRAM CONTROL STATEMENTS

```
        NAME
        DATA   132
        RDIF
        SDIF   12
        TRAN
        MA     ,4,1,1
        MA     ,6,12,2
        FST    30,,30
        SIDEN  ,,1
        FCST   ,30,132
```

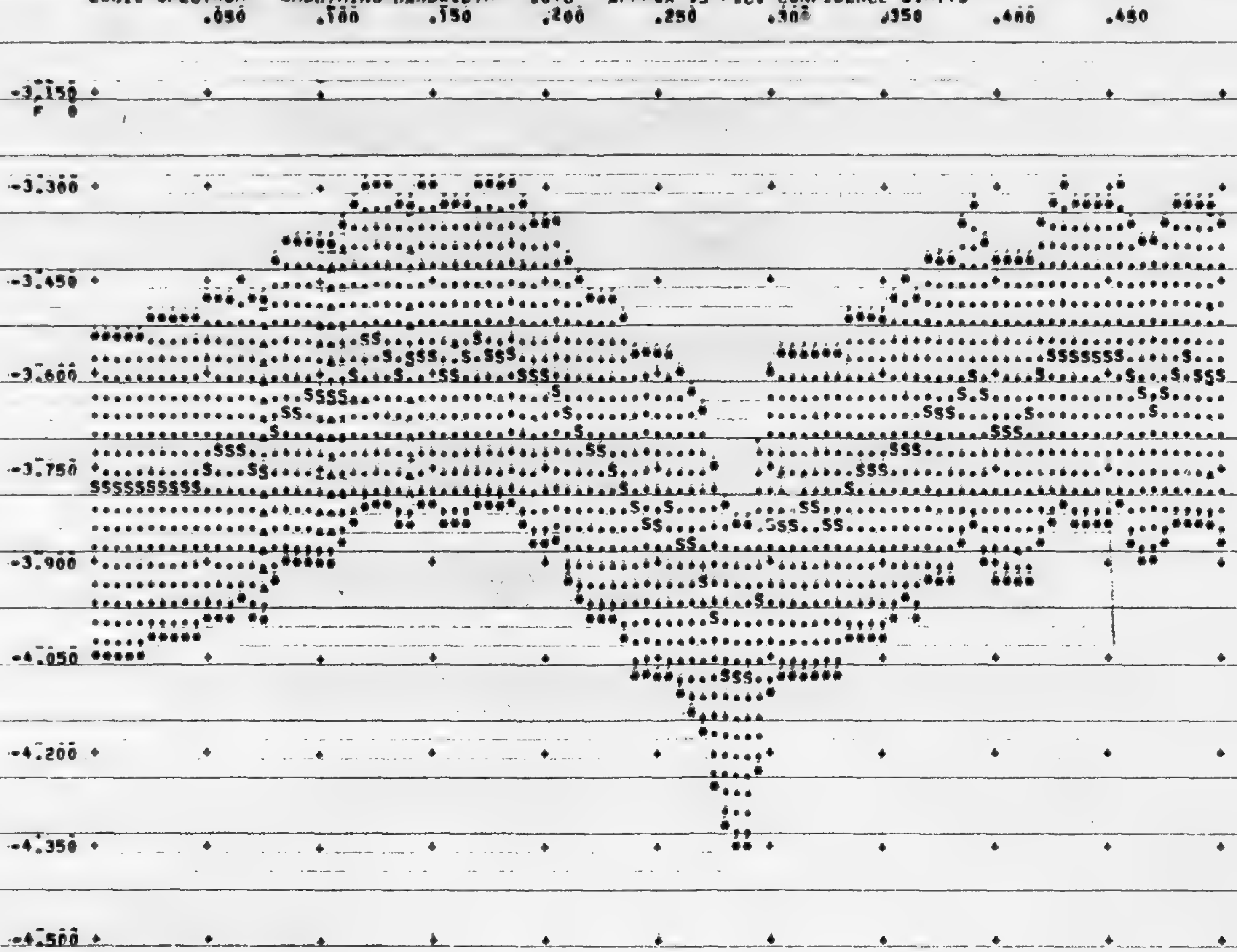
```
TIME SERIES PARAMETER ESTIMATION FOR MODEL  1
*******************************************************************************************************************
DATA - Z = INTERNATIONAL AIRLINE PASSENGERS - MONTHLY TOTALS - JAN 1949 - DEC 1960              132 OBSERVATIONS
DIFFERENCING ON Z - 1 OF ORDER  1, 1 OF ORDER 12
MODEL DEVELOPED WITH TRANSFORMED DATA = LOG(Z(T)+0.        )
*******************************************************************************************************************
PARAMETER              PARAMETER              PARAMETER        BEGINNING
 NUMBER                  TYPE                   ORDER            VALUE
*******************************************************************************************************************
    1               MOVING AVERAGE 1              1            .40000E+00
    2               MOVING AVERAGE 2             12            .60000E+00
*******************************************************************************************************************
EFFECTIVE NUMBER OF OBSERVATIONS FOR ESTIMATION IS   119
INITIAL SUM OF SQUARES =  .1674E+00


                                   ITERATION NO.   1

TEST POINT PARAMETER VALUES
   .3388E+00   .5691E+00

TEST POINT SUM OF SQUARES =  .1613E+00

PARAMETER VALUES VIA REGRESSION
      1           2
   .3388E+00   .5691E+00

SUM OF SQUARES AFTER REGRESSION =  .1612704E+00


                                   ITERATION NO.   2
```

```
TEST POINT PARAMETER VALUES

   .3289E+00   .5778E+00

TEST POINT SUM OF SQUARES =  .1612E+00

PARAMETER VALUES VIA REGRESSION

       1            2
   .3289E+00   .5778E+00

SUM OF SQUARES AFTER REGRESSION =  .1612339E+00

ITERATION STOPS - RELATIVE CHANGE IN SUM OF SQUARES LESS THAN  .1000E-02


CORRELATION MATRIX OF THE PARAMETERS

       1           2

 1  1.0000

 2  -.0662      1.0000

END OF ESTIMATION FOR MODEL  1
```

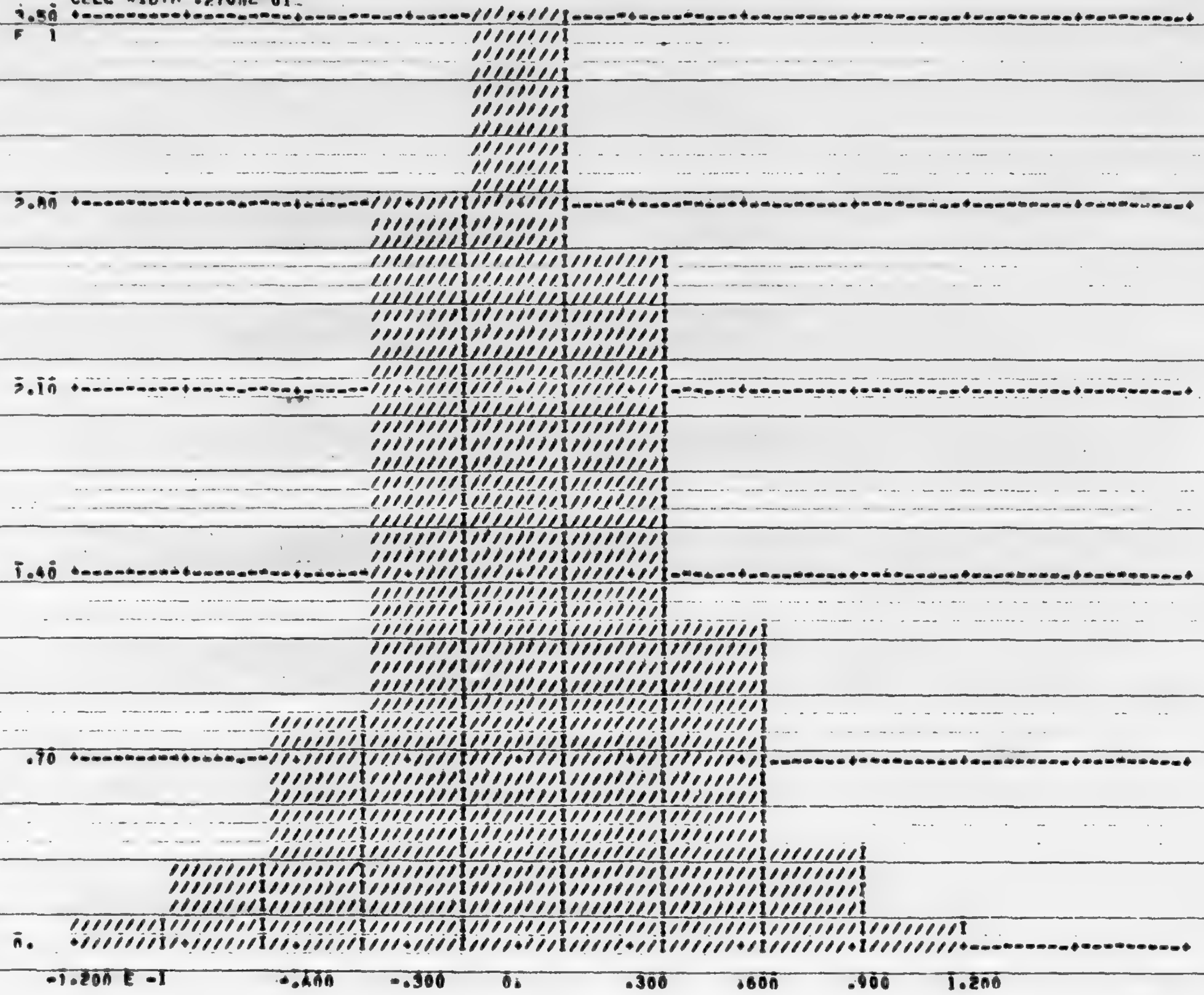
THE ESTIMATED RESIDUALS - MODEL 1
CUMULATIVE PERIODOGRAM .1 PROBABILITY LIMITS



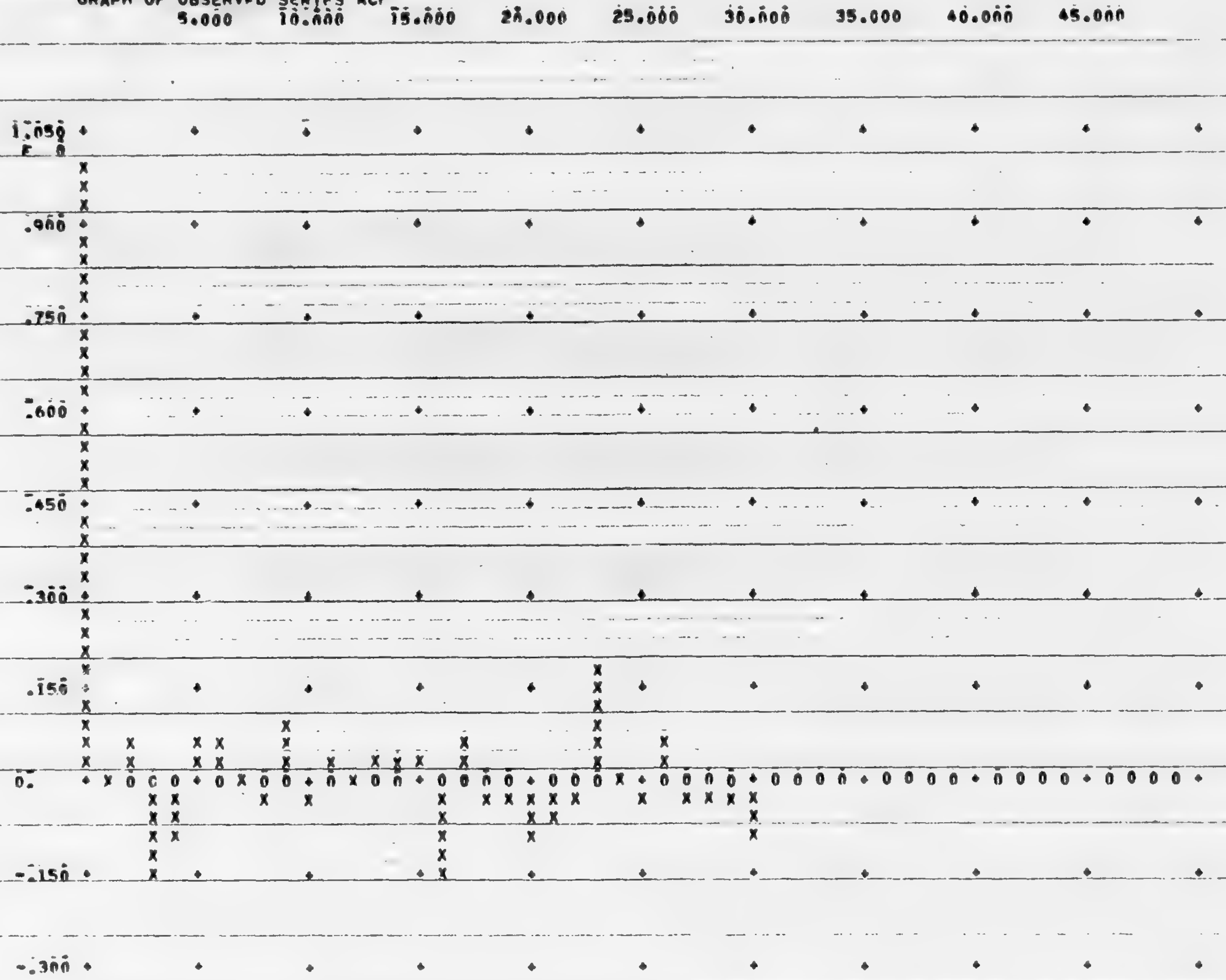
PREWHITENED INTERNATIONAL AIRLINE PASSENGERS - MONTHLY TOTALS - JAN 1949 - DEC 1
LOG10 SPECTRUM   SMOOTHING BANDWIDTH = .098   APPROX 95 P.C. CONFIDENCE LIMITS



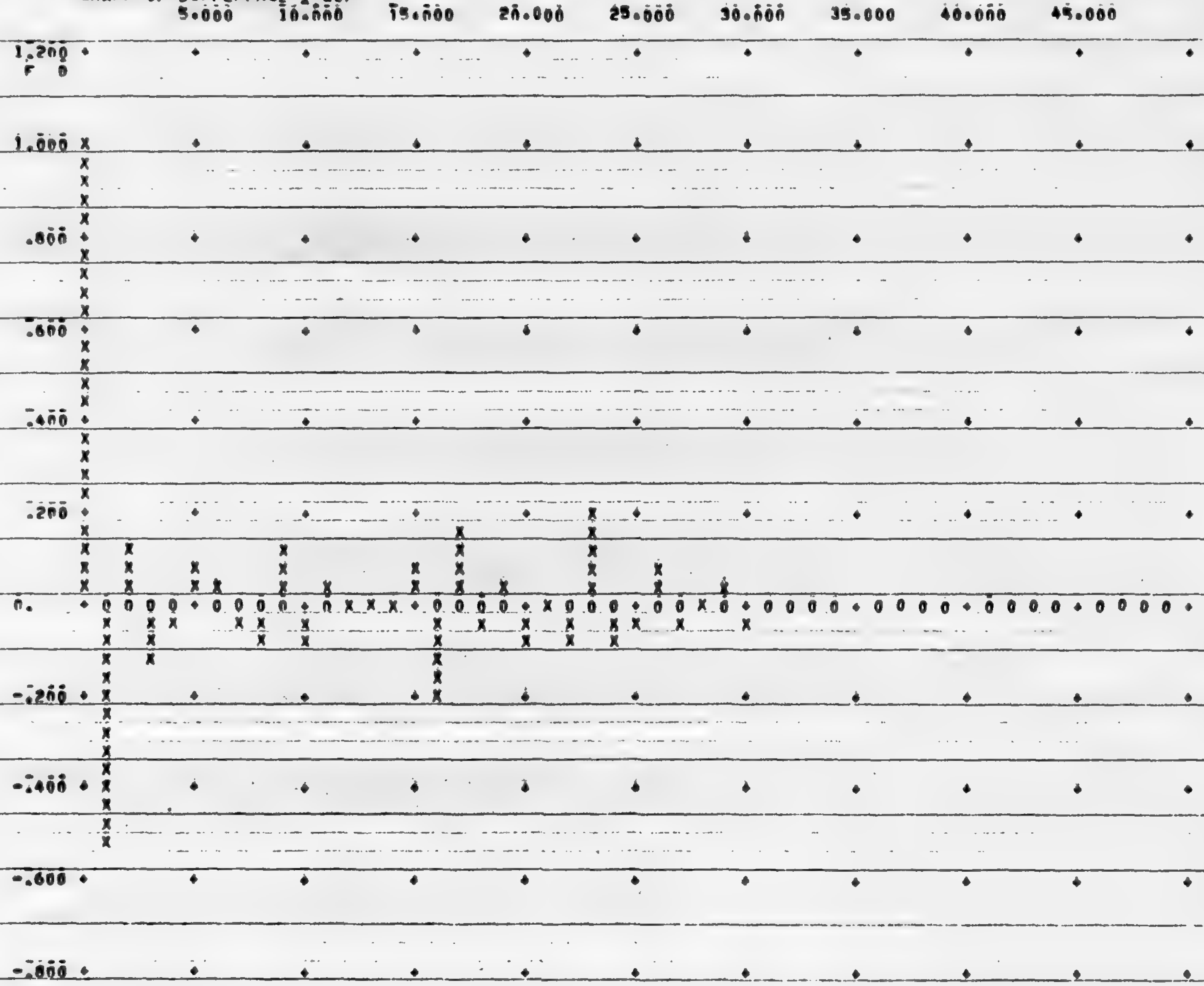
THE ESTIMATED RESIDUALS - MODEL 1
HISTOGRAM
CELL WIDTH .2700E-01



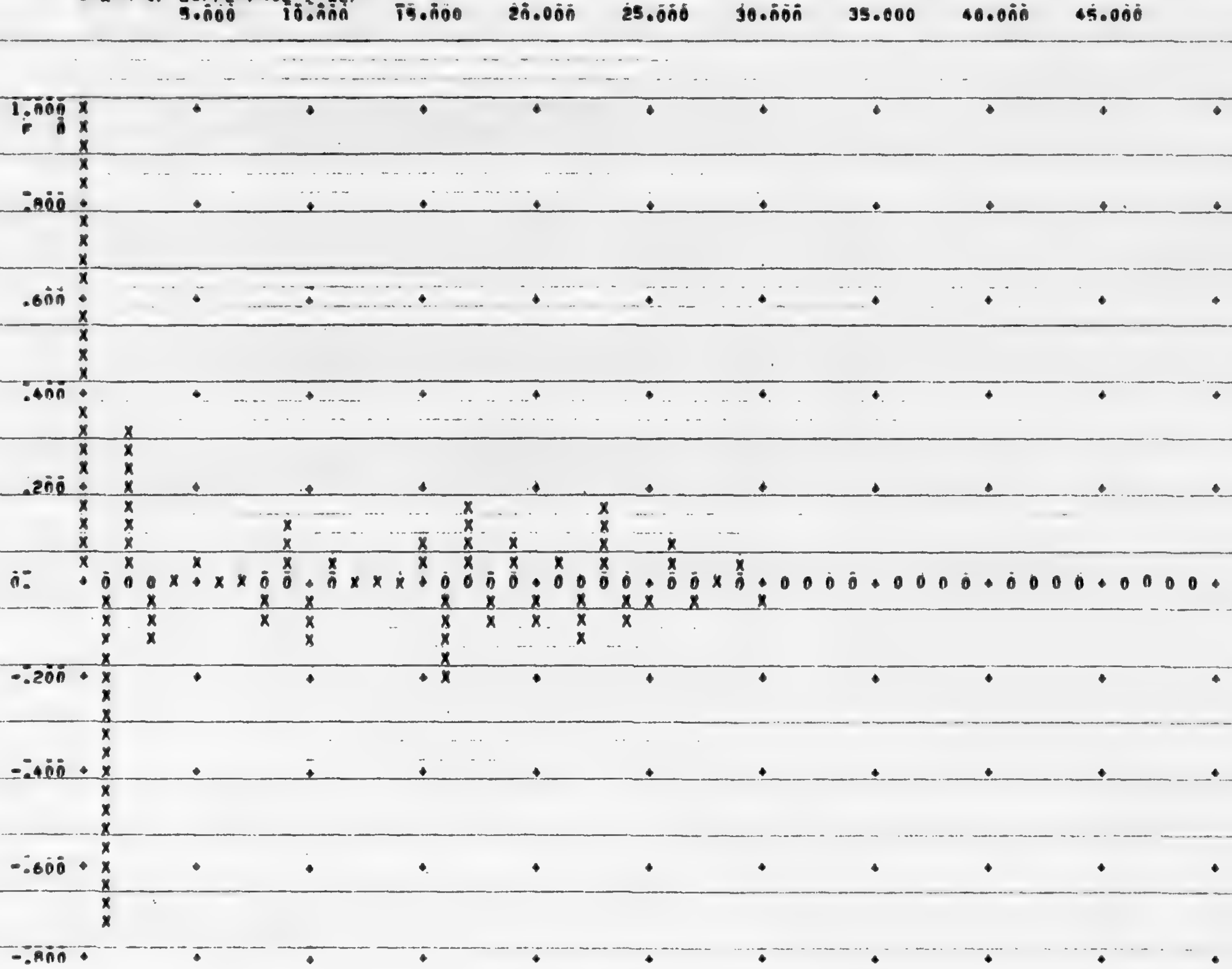
THE ESTIMATED RESIDUALS - MODEL 1
GRAPH OF OBSERVED SERIES ACF



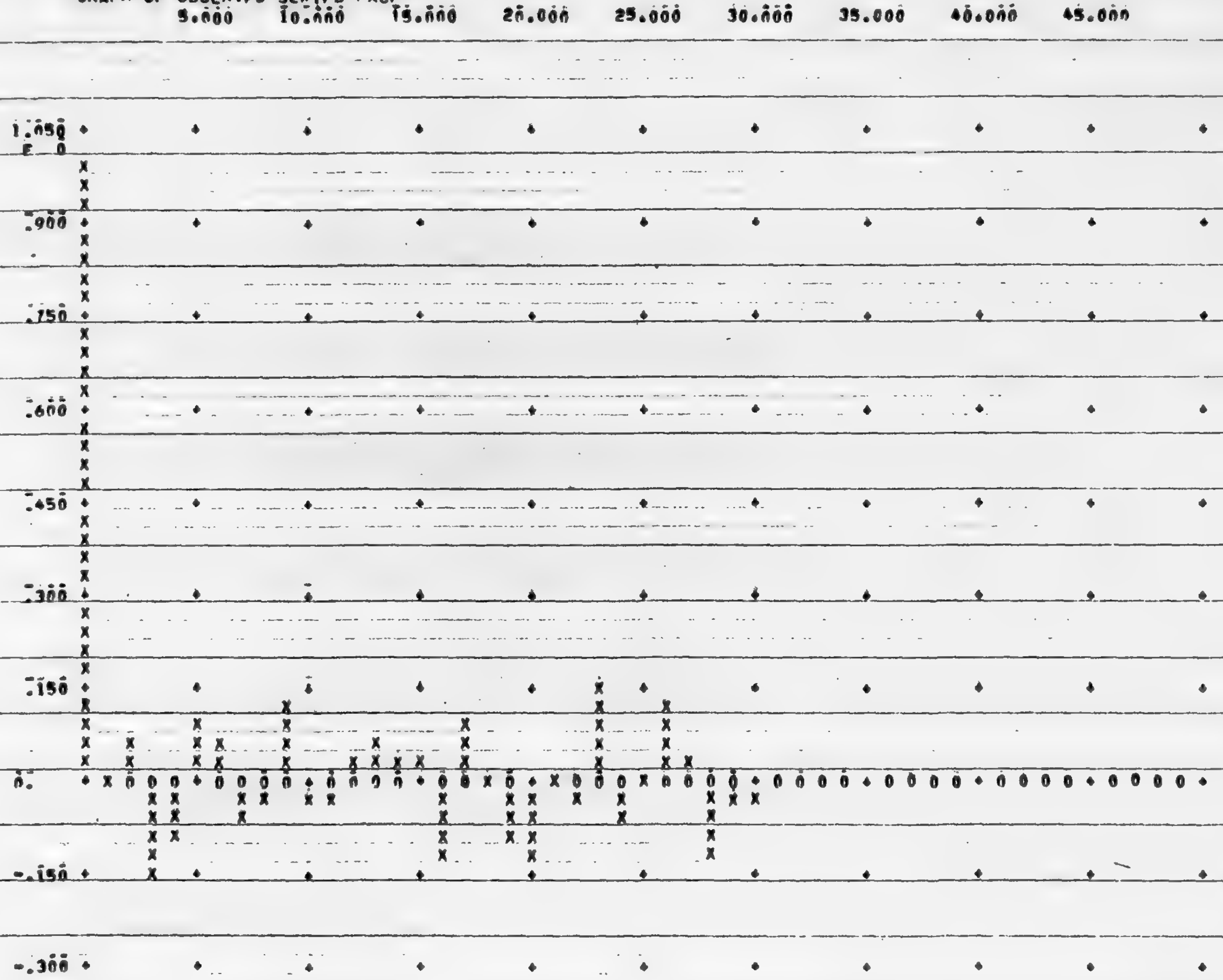
THE ESTIMATED RESIDUALS - MODEL 1
GRAPH OF DIFFERENCE 1 ACF



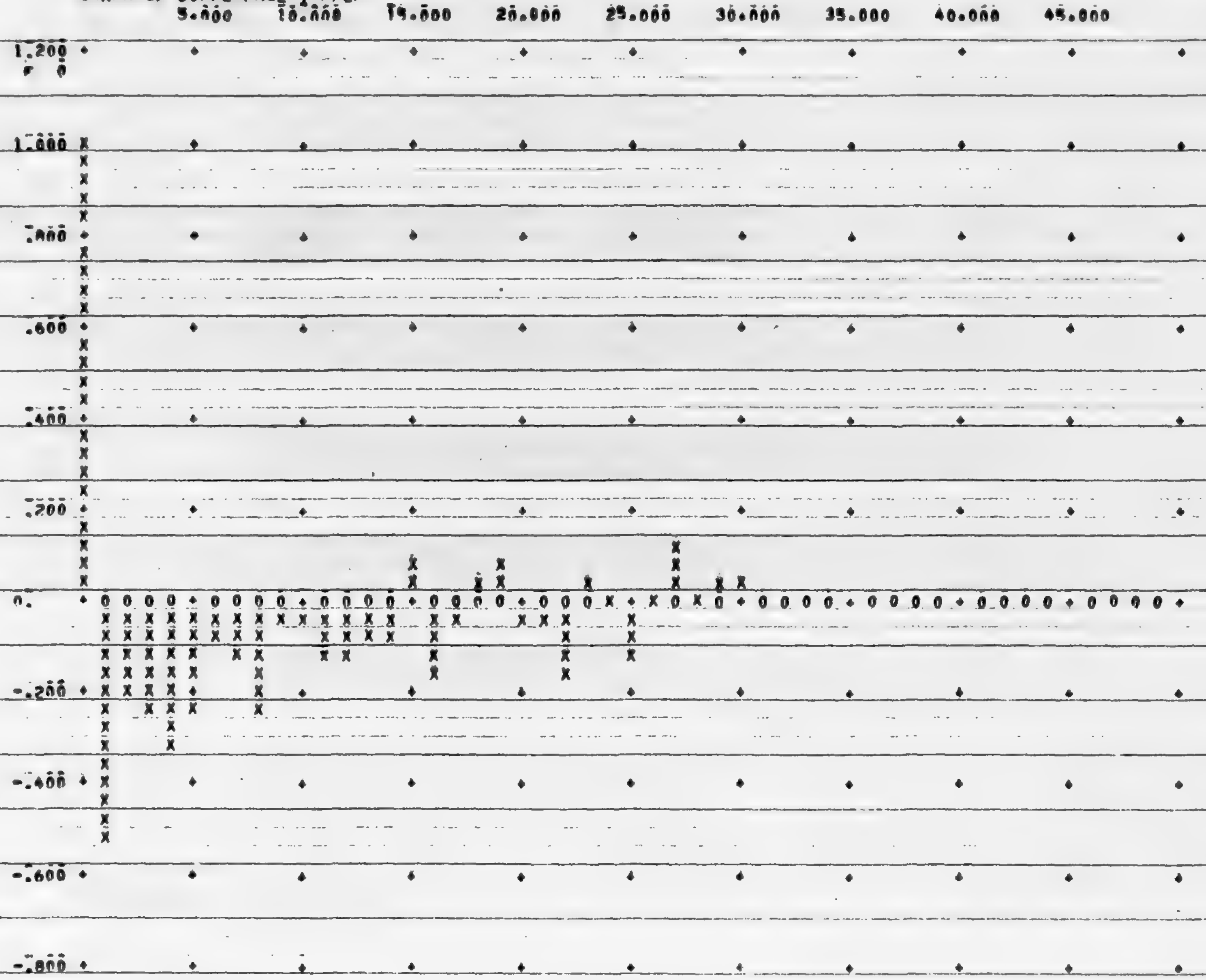
THE ESTIMATED RESIDUALS - MODEL 1
GRAPH OF DIFFERENCE 2 ACF



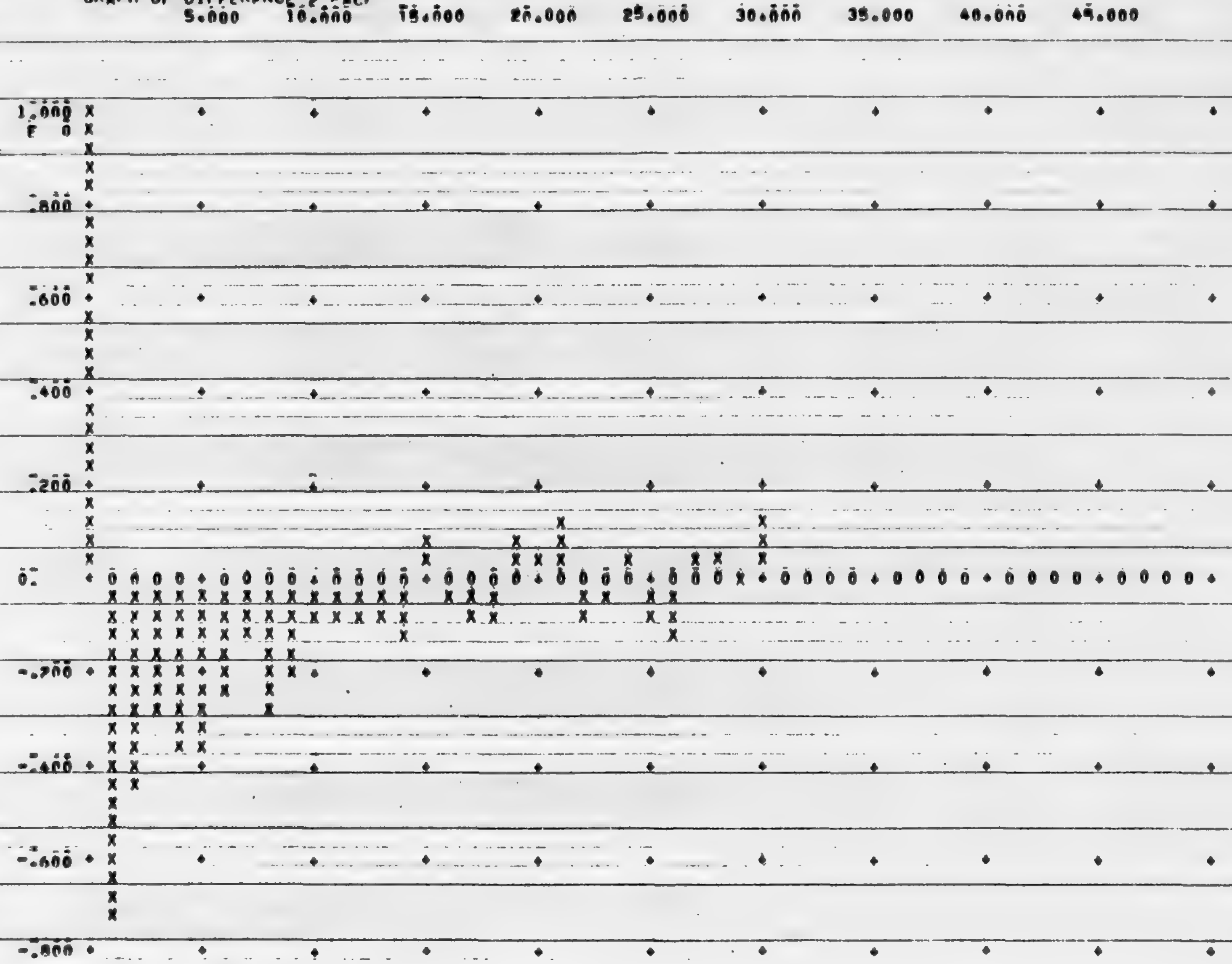
THE ESTIMATED RESIDUALS - MODEL 1
GRAPH OF OBSERVED SERIES PACF



THE ESTIMATED RESIDUALS - MODEL 1
GRAPH OF DIFFERENCE 1 PACF



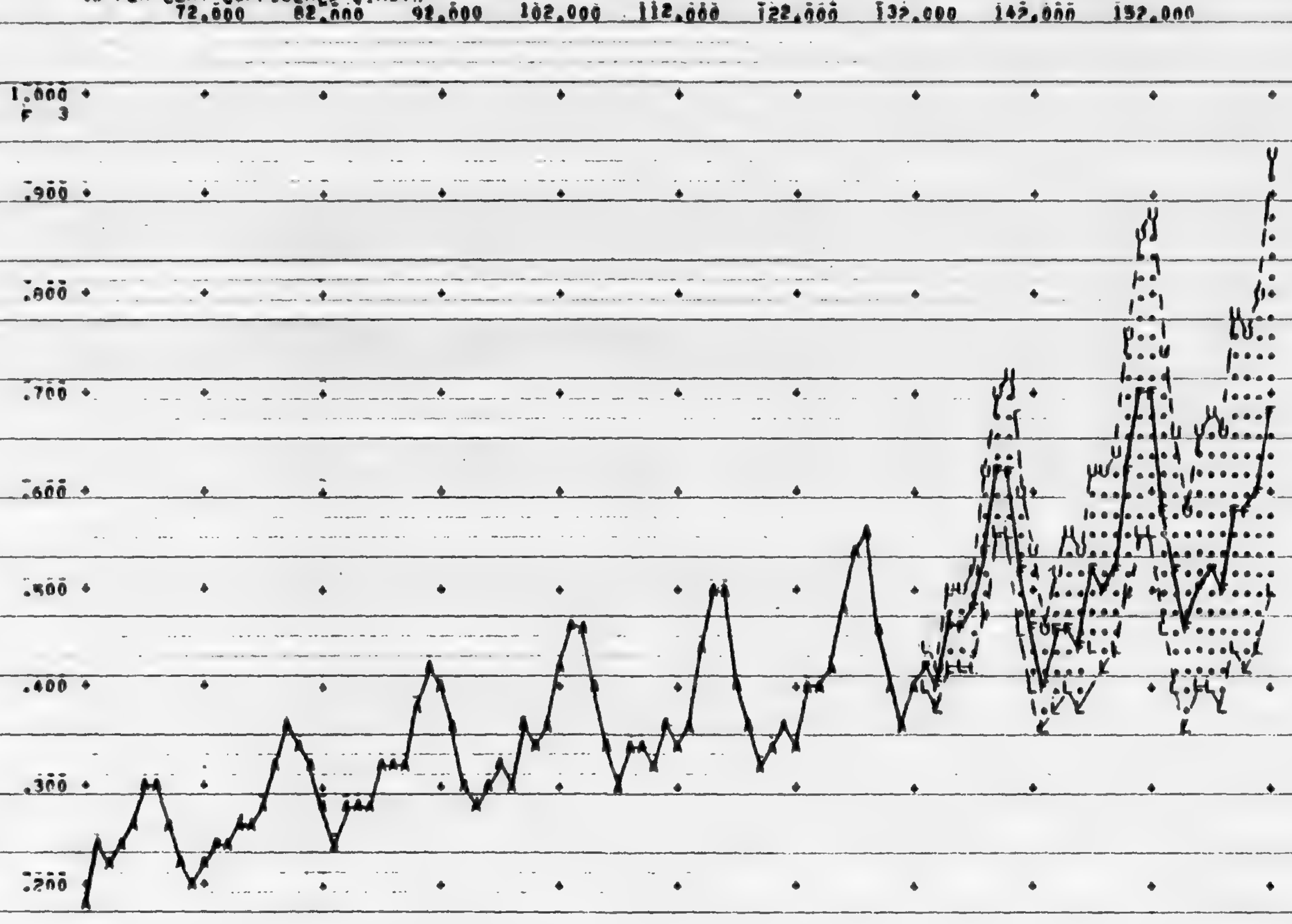
THE ESTIMATED RESIDUALS - MODEL 1
GRAPH OF DIFFERENCE 2 PACF



```
SUMMARY OF MODEL  1
************************************************************************************************************
DATA - Z = INTERNATIONAL AIRLINE PASSENGERS - MONTHLY TOTALS - JAN 1949 - DEC 1960               132 OBSERVATIONS
DIFFERENCING ON Z - 1 OF ORDER  1, 1 OF ORDER 12
MODEL DEVELOPED WITH TRANSFORMED DATA = LOG(Z(T)+0.        )
************************************************************************************************************
PARAMETER           PARAMETER              PARAMETER       ESTIMATED                  95 PER CENT
 NUMBER               TYPE                   ORDER           VALUE          LOWER LIMIT            UPPER LIMIT
************************************************************************************************************
    1              MOVING AVERAGE 1            1          .32889E+00         .15514E+00            .50265E+00
    2              MOVING AVERAGE 2           12          .57784E+00         .41813E+00            .73756E+00
************************************************************************************************************
OTHER INFORMATION AND RESULTS
************************************************************************************************************
RESIDUAL SUM OF SQUARES       .16123E+00    117 D.F.            RESIDUAL MEAN SQUARE              .13781E-02
NUMBER OF RESIDUALS                  119                        RESIDUAL STANDARD ERROR           .37122E-01
```

```
TIME SERIES FORECASTING FOR MODEL  1
********************************************************************************************************************
DATA - Z = INTERNATIONAL AIRLINE PASSENGERS - MONTHLY TOTALS - JAN 1949 - DEC 1960               132 OBSERVATIONS
DIFFERENCING ON Z - 1 OF ORDER  1, 1 OF ORDER 12
MODEL DEVELOPED WITH TRANSFORMED DATA = LOG(Z(T)+0.         )
********************************************************************************************************************
PARAMETER               PARAMETER              PARAMETER        ESTIMATED
 NUMBER                   TYPE                   ORDER            VALUE
********************************************************************************************************************
    1                MOVING AVERAGE 1              1          .32889E+00
    2                MOVING AVERAGE 2             12          .57784E+00
********************************************************************************************************************
NUMBER OF TIME ORIGINS FOR FORECASTS =            1
NUMBER OF FORECASTS AT EACH TIME ORIGIN =        30
FORECAST TIME ORIGINS ARE T=                    132
```

INTERNATIONAL AIRLINE PASSENGERS - MONTHLY TOTALS - JAN 1949 - DEC 1960
FORECAST AT ORIGIN 132   F=FORECAST A=ACTUAL
90 PER CENT CONFIDENCE LIMITS



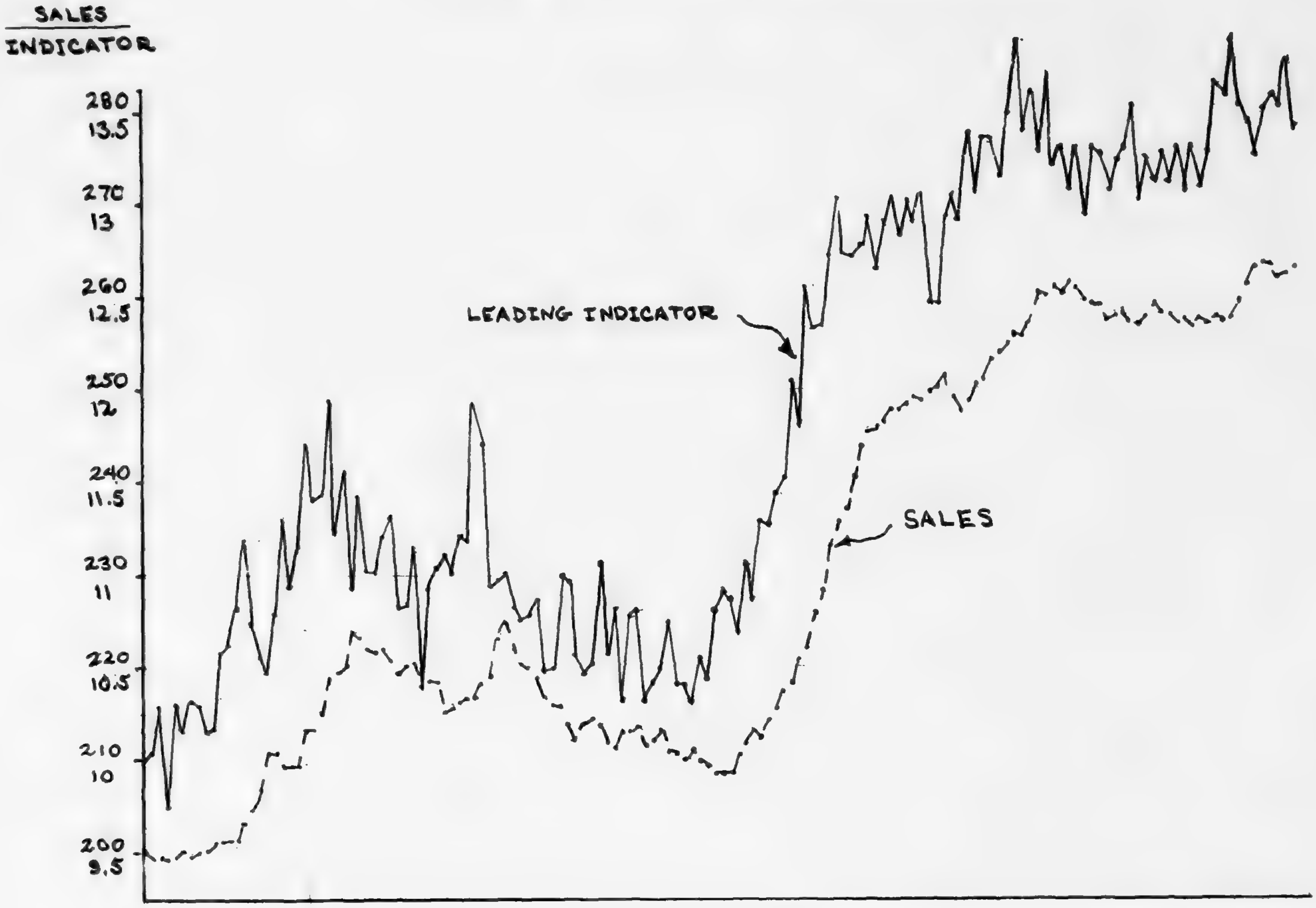
```
WEIGHTS USED IN CALCULATING CONFIDENCE LIMITS
AND UPDATING FORECASTS AFTER NEW OBSERVATION

         J                  PSI(J)

         0              .1000000E+01
         1              .6711056E+00
         2              .6711056E+00
         3              .6711056E+00
         4              .6711056E+00
         5              .6711056E+00
         6              .6711056E+00
         7              .6711056E+00
         8              .6711056E+00
         9              .6711056E+00
        10              .6711056E+00
        11              .6711056E+00
        12              .1093261E+01
        13              .9544164E+00
        14              .9544164E+00
        15              .9544164E+00
        16              .9544164E+00
        17              .9544164E+00
        18              .9544164E+00
        19              .9544164E+00
        20              .9544164E+00
        21              .9544164E+00
        22              .9544164E+00
        23              .9544164E+00
        24              .1376572E+01
        25              .1237727E+01
        26              .1237727E+01
        27              .1237727E+01
        28              .1237727E+01
        29              .1237727E+01
        30              .1237727E+01
```

APPENDIX III

An Example of Transfer Function Modelling and Forecasting

Data for the following example is given in Reference [1], p. 537. A transfer function model for this data is presented in Section 11.5.3 of Reference [1].

SALES DATA WITH LEADING INDICATOR

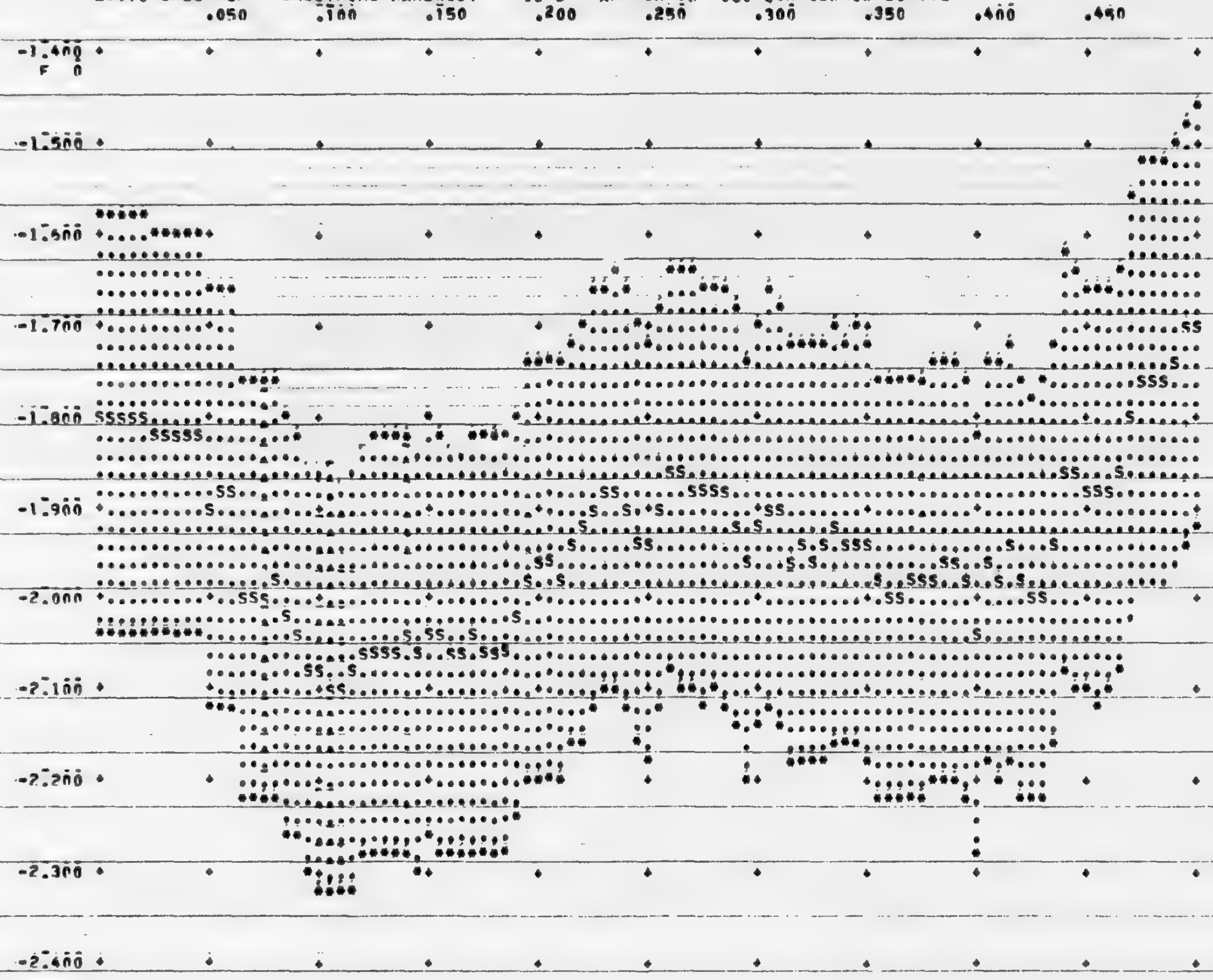


PROGRAM CONTROL STATEMENTS

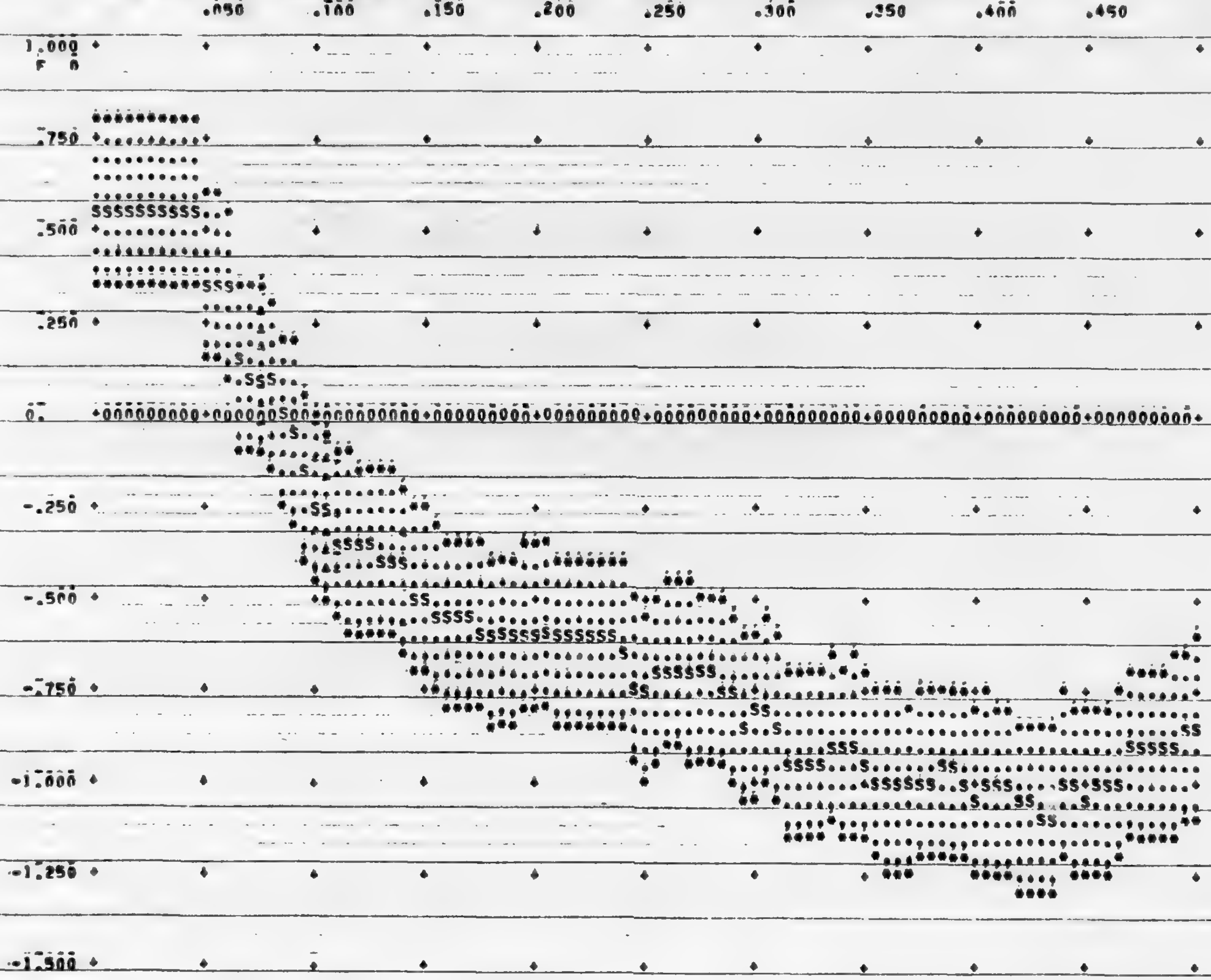
```
        NAME
        DATA  150
        YNAME
        YDATA 150
        YRDIF
        RDIF
        MA    .453,1,1
        TIDEA 35,,15,35,,35
        SIDEA ,,1
```

```
SPECIFICATION OF THE PREWHITENING TRANSFORMATION APPLIED TO INPUT AND OUTPUT SERIES
*****************************************************************************************************

DATA - X = LEADING INDICATOR                                                          150 OBSERVATIONS
       Y = SALES DATA

DIFFERENCING ON Y - 1 OF ORDER  1

DIFFERENCING ON X - 1 OF ORDER  1

*****************************************************************************************************

PARAMETER             PARAMETER              PARAMETER        ESTIMATED
 NUMBER                 TYPE                   ORDER            VALUE

*****************************************************************************************************

    1               MOVING AVERAGE 1               1            .45300E+00

*****************************************************************************************************
```

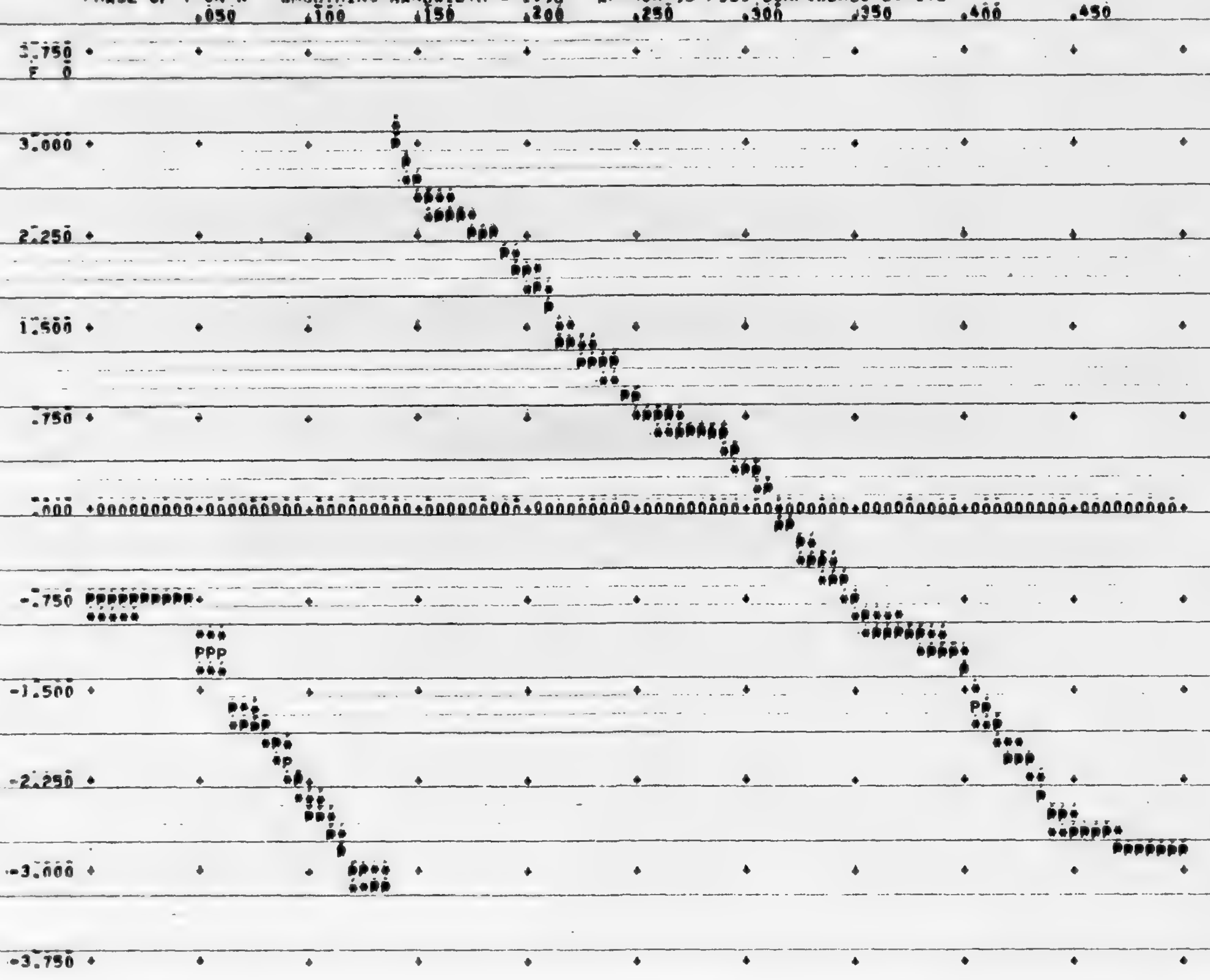
PREWHITENED LEADING INDICATOR
LOG10 SPECTRUM   SMOOTHING BANDWIDTH = .098   APPROX 95 P.C. CONFIDENCE LIMITS



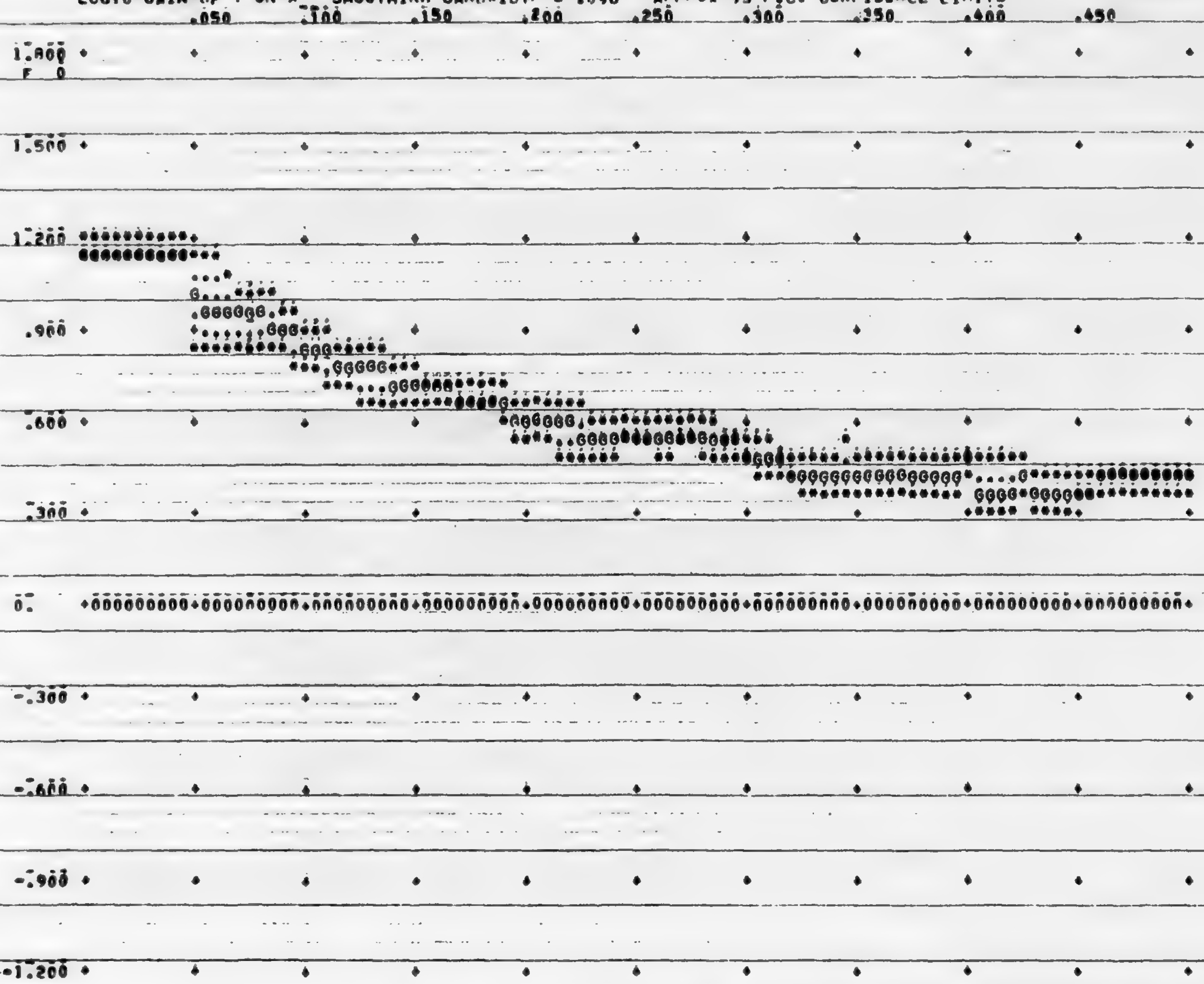
PREWHITENED LEADING INDICATOR
LOG10 SPECTRUM SHOOTHING BANDWIDTH = .098 APPROX 95 P.C. CONFIDENCE LIMITS



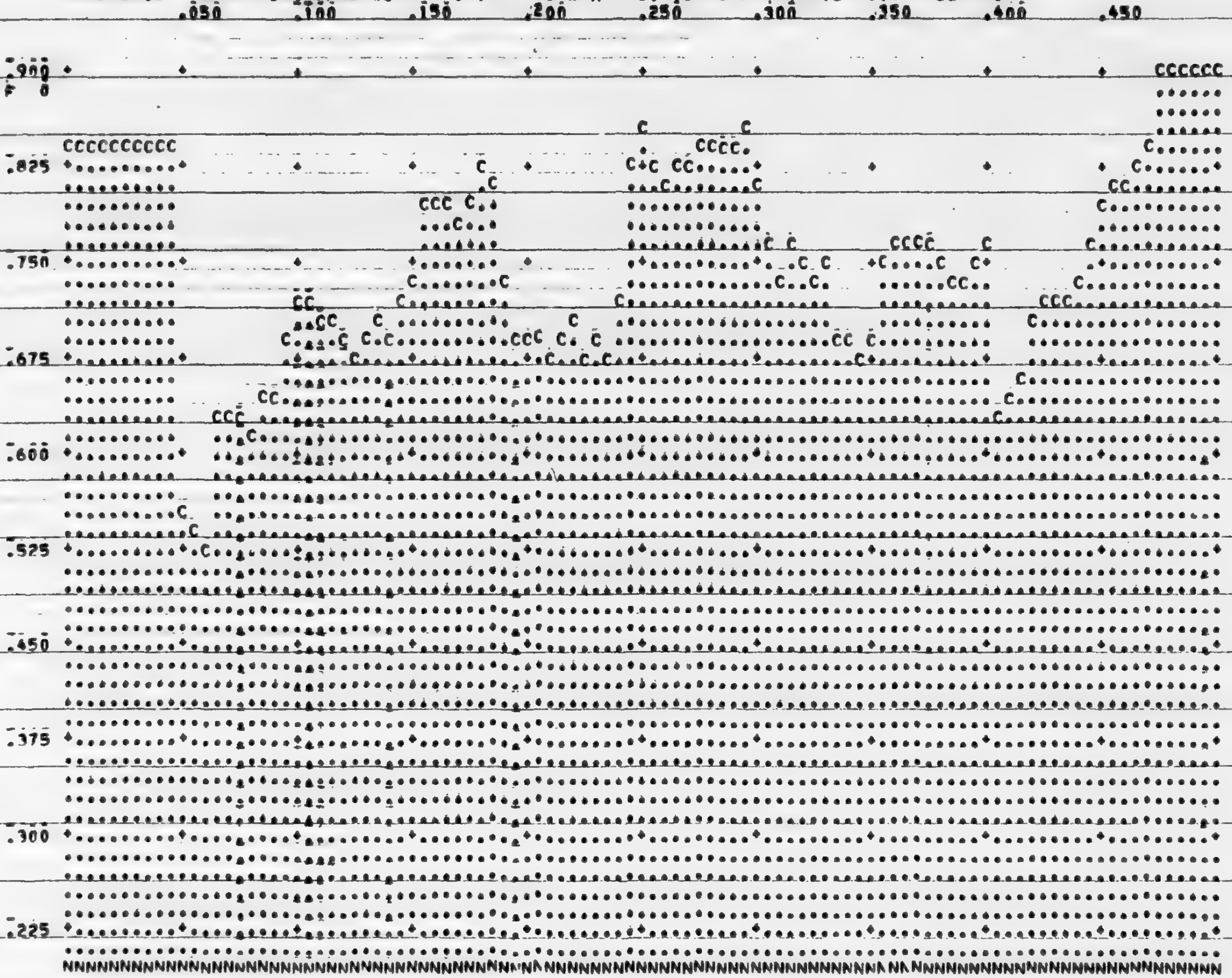
```
X = PREWHITENED LEADING INDICATOR
Y = PREWHITENED SALES DATA
PHASE OF Y ON X   SMOOTHING BANDWIDTH = .098   APPROX 95 P.C. CONFIDENCE LIMITS
```



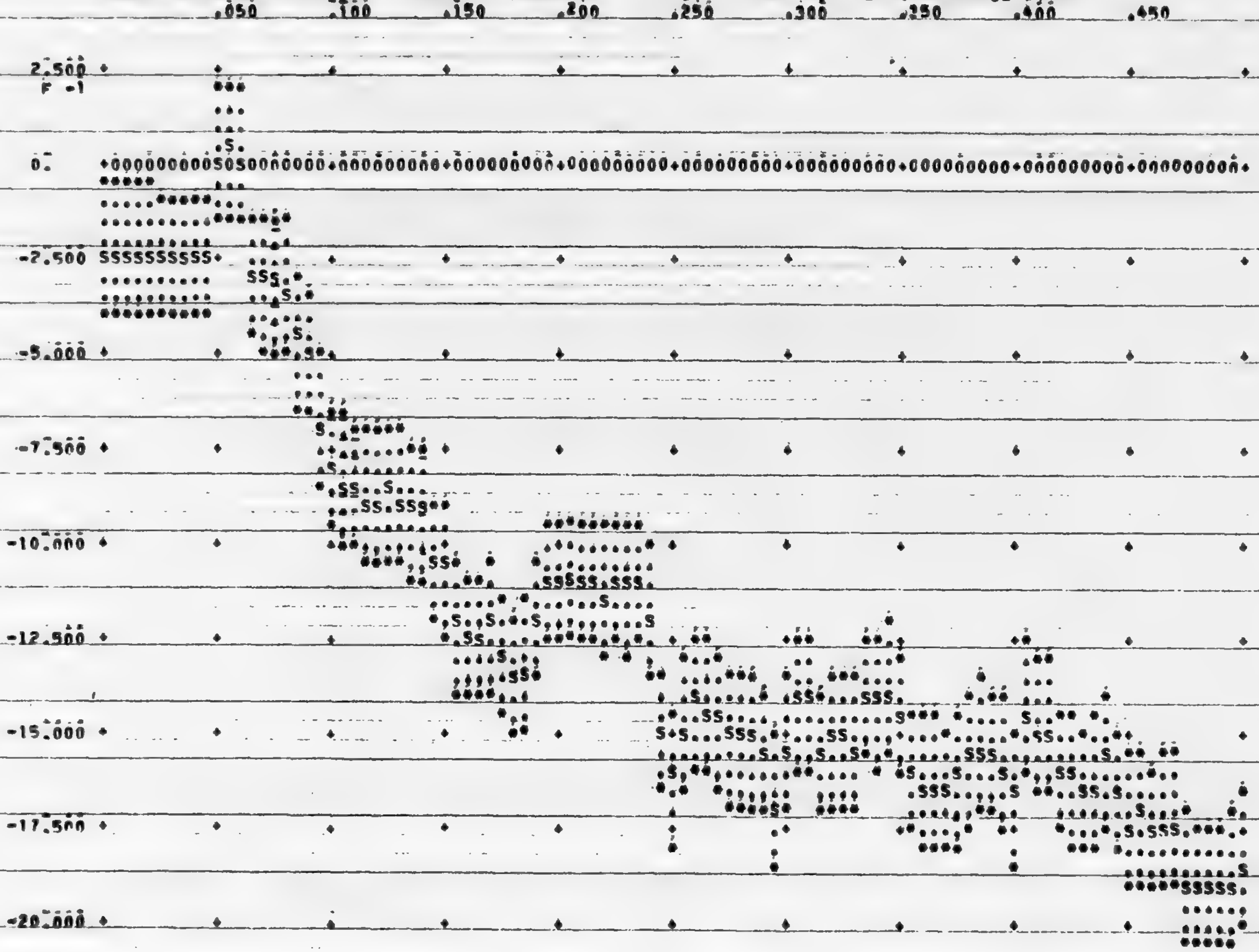
X - PREWHITENED LEADING INDICATOR
Y - PREWHITENED SALES DATA
LOGIO GAIN OF Y ON X   SMOOTHING BANDWIDTH = .098   APPROX 95 P.C. CONFIDENCE LIMITS



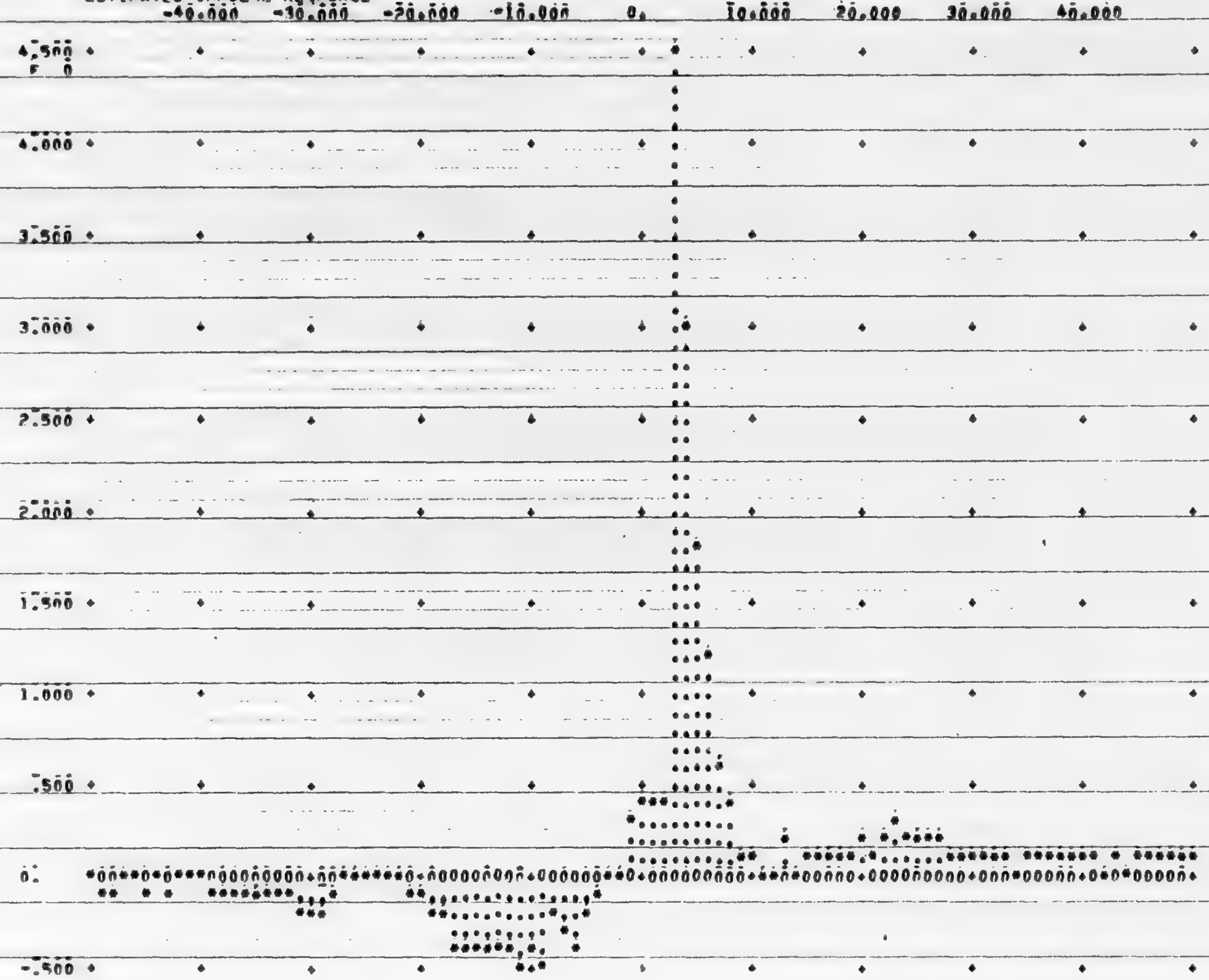
```
X - PREWHITENED LEADING INDICATOR
Y - PREWHITENED SALES DATA
COHERENCE   SMOOTHING BANDWIDTH = .098   LINE AT .19 GIVES APPROX 95 P.C. NULL POINT
```



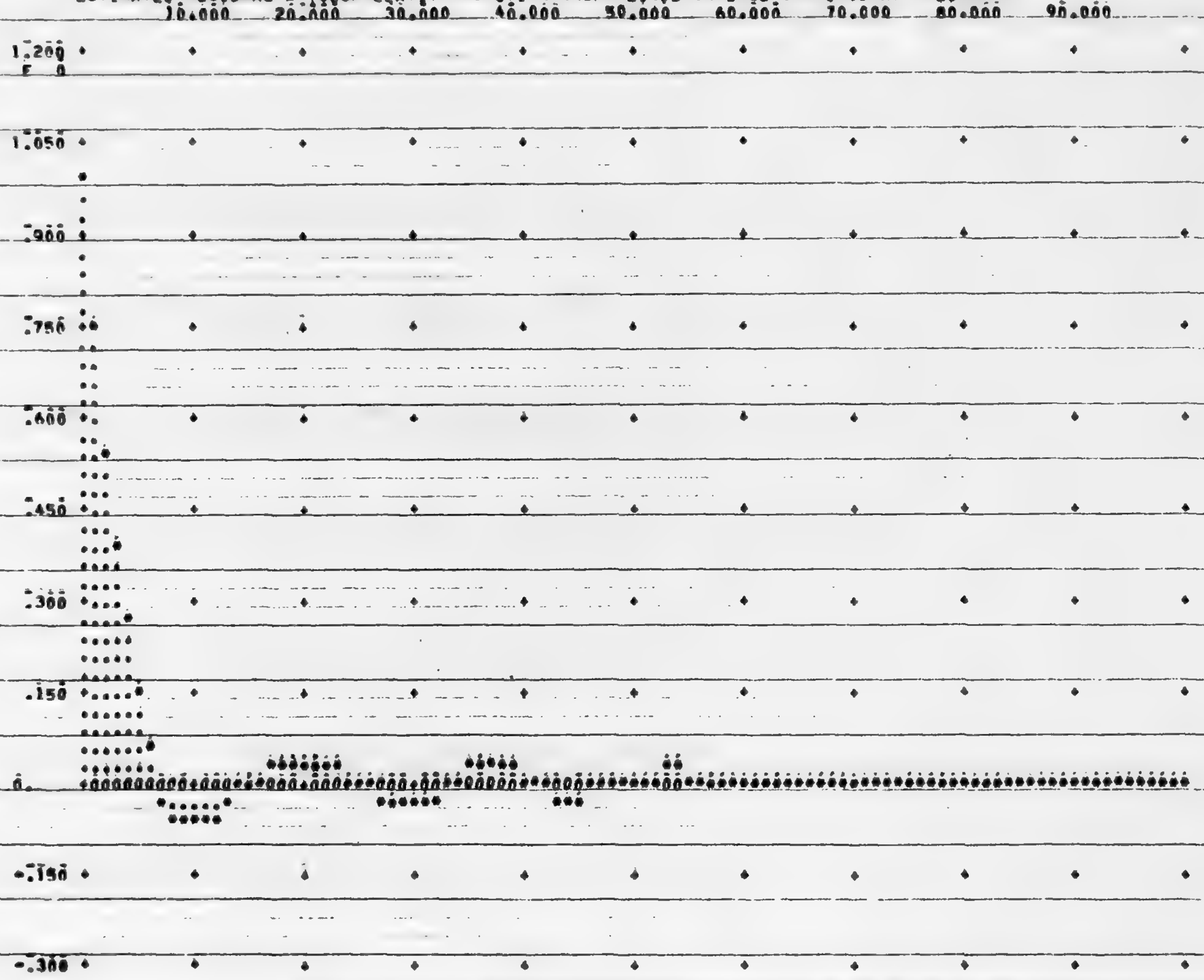
X - PREWHITENED LEADING INDICATOR
Y - PREWHITENED SALES DATA
LOG10 RESIDUAL SPECTRUM SMOOTHING BANDWIDTH = .098 APPROX 95 P.C. CONFIDENCE LIMITS



X - PREWHITENED LEADING INDICATOR
Y - PREWHITENED SALES DATA
ESTIMATED IMPULSE RESPONSE

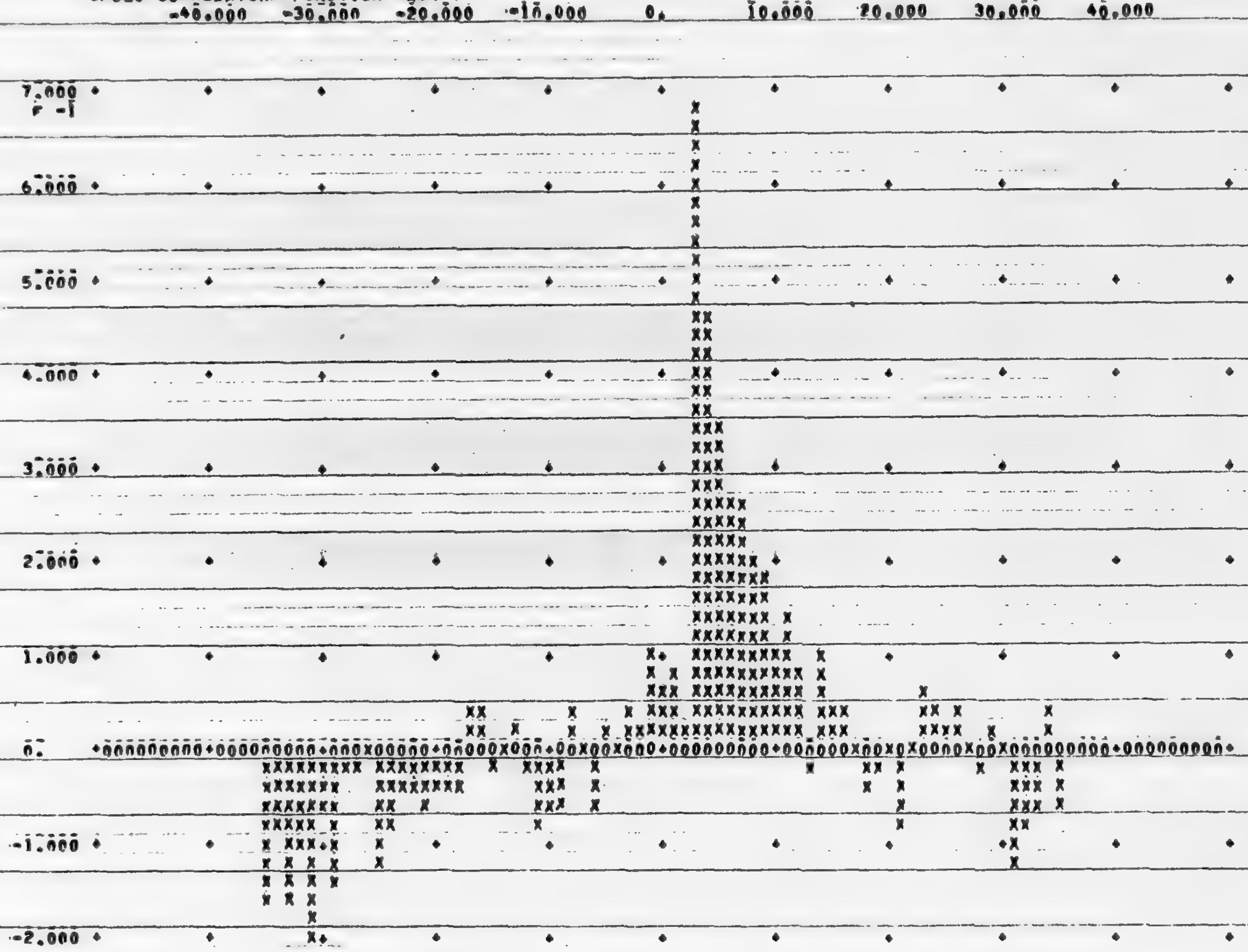


X - PREWHITENED LEADING INDICATOR
Y - PREWHITENED SALES DATA
ESTIMATED RESIDUAL AUTOCORRELATION FUNCTION (NORMALIZED INVERSE OF RESIDUAL SPECTRUM)



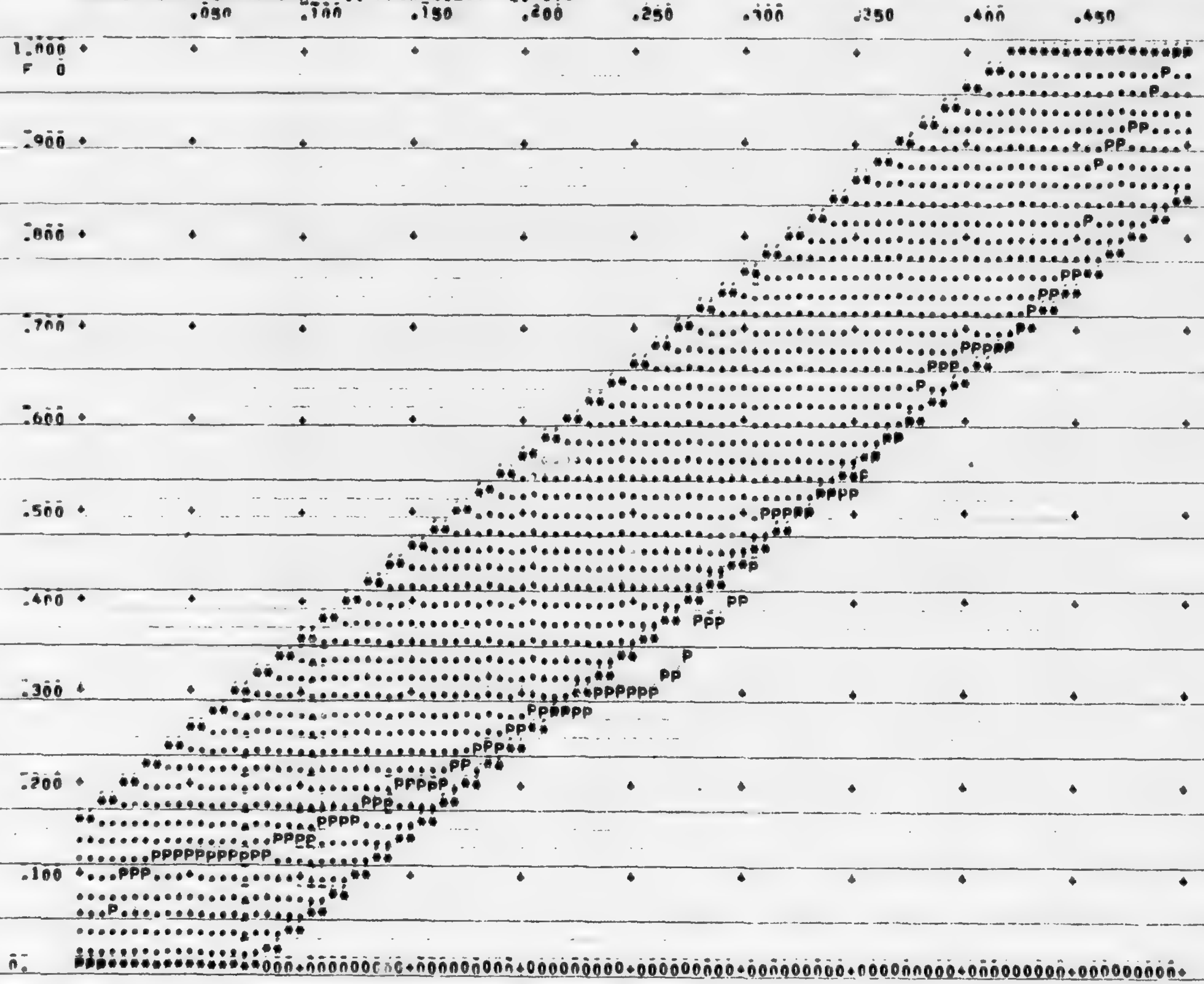
```
SPECIFICATION OF THE PREWHITENING TRANSFORMATION APPLIED TO INPUT AND OUTPUT SERIES

*******************************************************************************************************

DATA - X = LEADING INDICATOR                                                               150 OBSEPVATIONS
       Y = SALES DATA

DIFFERENCING ON Y - 1 OF ORDER  1

DIFFERENCING ON X - 1 OF ORDER  1

*******************************************************************************************************

PARAMETER          PARAMETER          PARAMETER      ESTIMATED
 NUMBER              TYPE               ORDER          VALUE

*******************************************************************************************************

    1            MOVING AVERAGE 1         1          .45300E+00

*******************************************************************************************************

EFFECTIVE NUMBER OF OBSERVATIONS FOR CROSS CORRELATION CALCULATION IS  149
```

A - PREWHITENED LEADING INDICATOR
B - PREWHITENED SALES DATA
CROSS CORRELATION FUNCTION RAB(K)

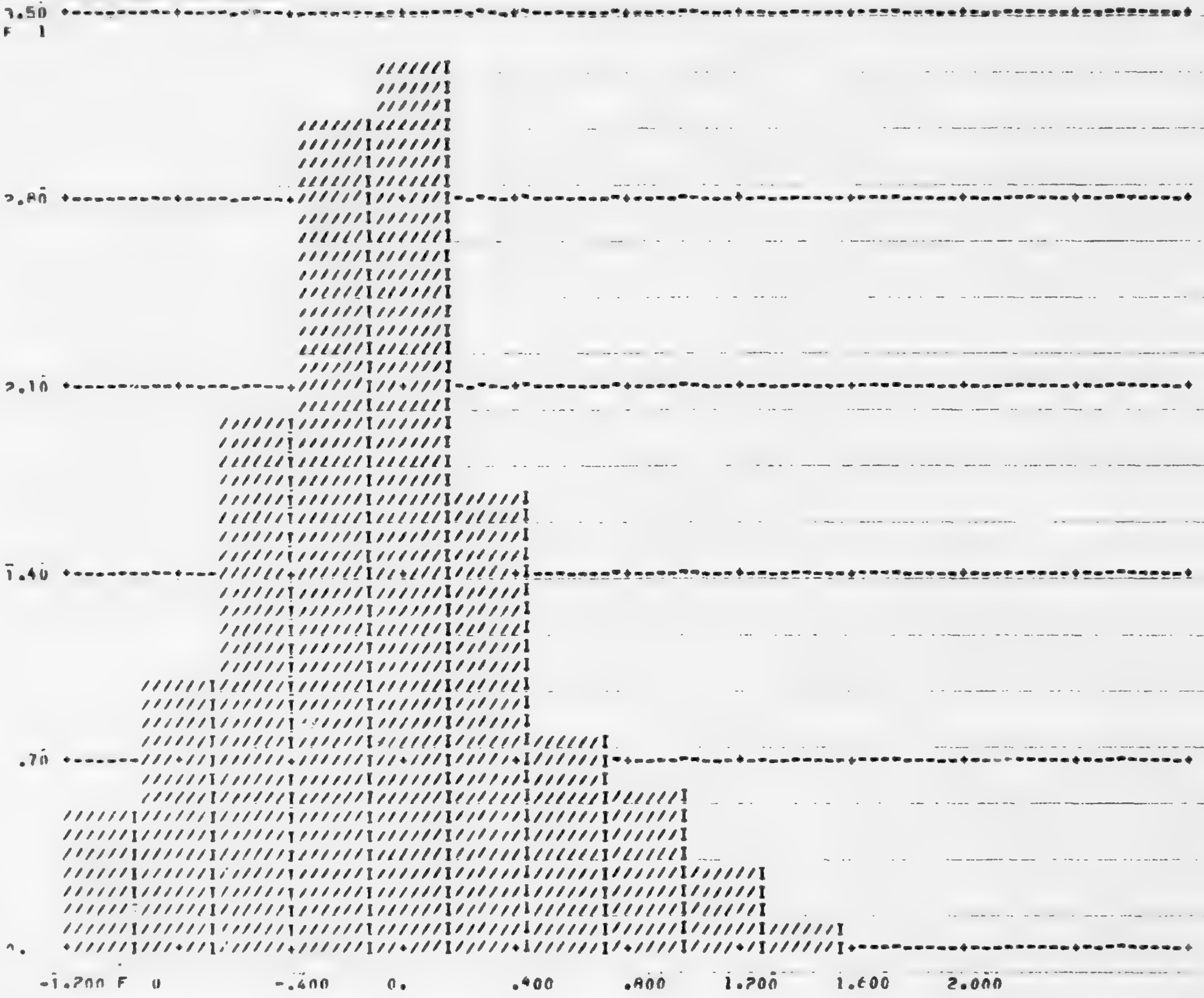


```
ESTIMATED IMPULSE RESPONSE WEIGHTS V(K)

        K              V(K)

        0              .451
        1              .570
        2              .162
        3             4.734
        4             3.190
        5             2.413
        6             1.826
        7             1.894
        8             1.402
        9             1.225
       10              .496
       11             1.019
       12              .491
       13             -.097
       14              .757
       15              .329
       16              .305
       17             -.051
       18             -.269
       19             -.175
       20             -.030
       21             -.555
       22             -.048
       23              .481
       24              .227
       25              .150
       26              .237
       27             -.065
       28             -.086
       29              .103
       30              .024
       31             -.852
       32             -.573
       33             -.484
       34              .250
       35             -.434
```

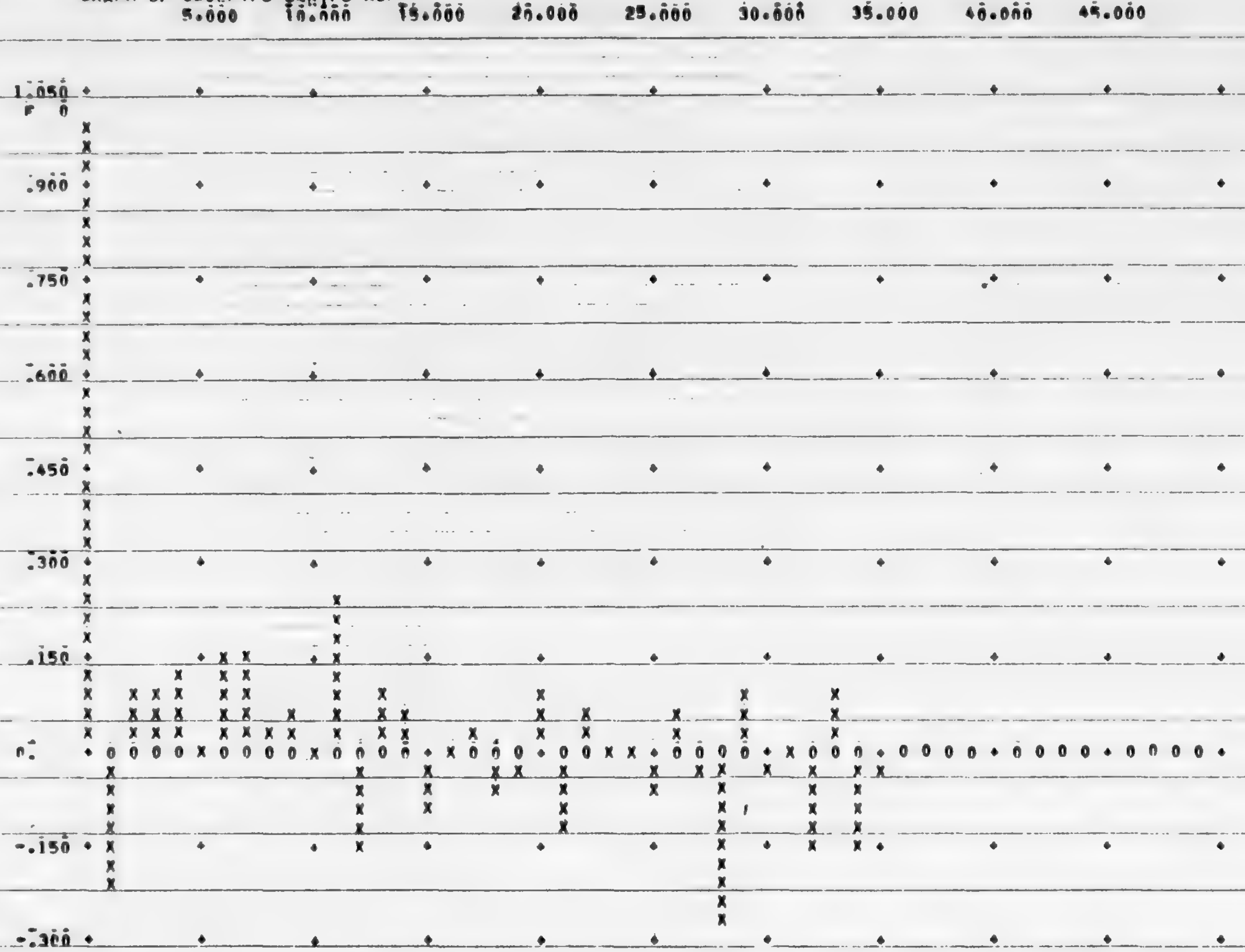
THE GENERATED NOISE SERIES
CUMULATIVE PERIODOGRAM .1 PROBABILITY LIMITS



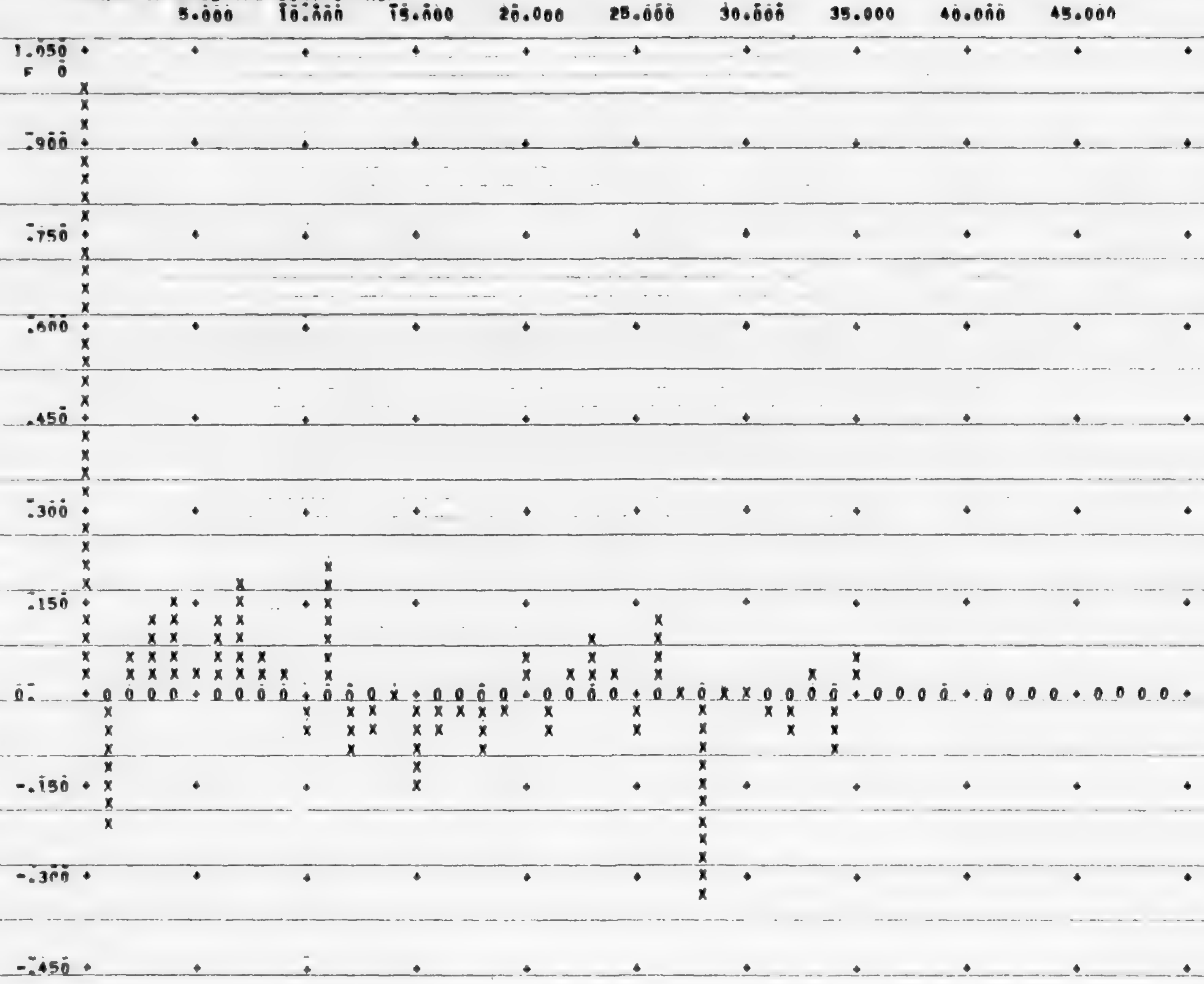
THE GENERATED NOISE SERIES
HISTOGRAM
CELL WIDTH .2800E+00



THE GENERATED NOISE SERIES
GRAPH OF OBSERVED SERIES ACF



THE GENERATED NOISE SERIES
GRAPH OF OBSERVED SERIES PACF



PROGRAM CONTROL STATEMENTS

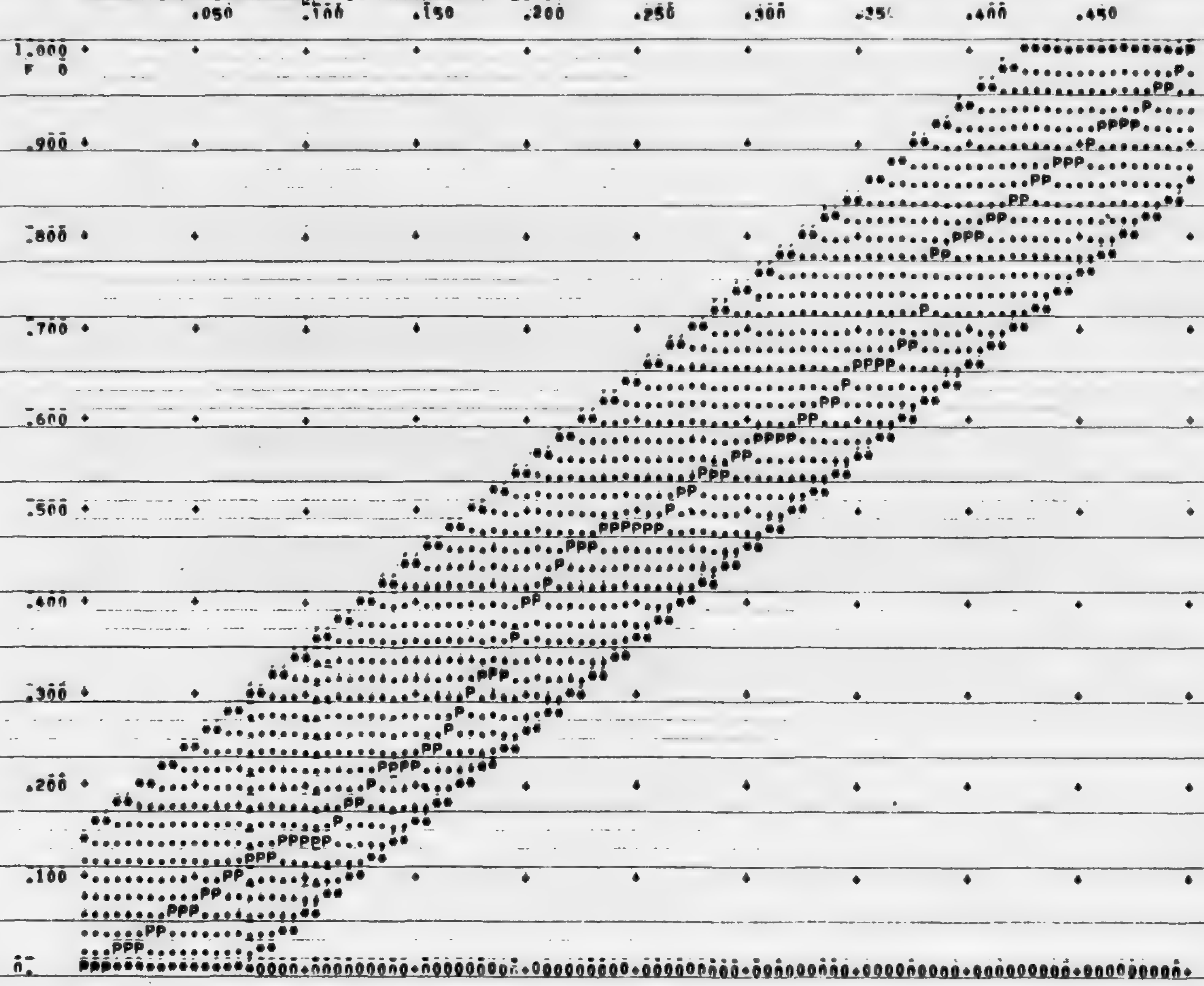
```
        NAME
        DATA  150
        YNAME
        YDATA 150
        YRDIF
        RDIF
        MA    .453,1,1
        NMA   .5,1,1
        TLAG  4,8,0,1
        OLAG  .7,1,1
        P     3
        TEST  35,35,.35
```

```
TRANSFER FUNCTION MODEL ESTIMATION FOR MODEL 1

****************************************************************************************************

DATA - X = LEADING INDICATOR                                                      150 OBSERVATIONS
       Y = SALES DATA

DIFFERENCING ON Y - 1 OF ORDER  1

DIFFERENCING ON X - 1 OF ORDER  1

****************************************************************************************************

NOISE MODEL PARAMETERS

****************************************************************************************************

PARAMETER           PARAMETER             PARAMETER       BEGINNING
 NUMBER               TYPE                  ORDER           VALUE

****************************************************************************************************

    1               MOVING AVERAGE 1          1           .50000E+00

****************************************************************************************************

TRANSFER FUNCTION PARAMETERS

****************************************************************************************************

    2               OUTPUT LAG 1              1           .70000E+00

    3               INPUT LAG 1               0           .40000E+01

****************************************************************************************************

****************************************************************************************************

ESTIMATION FOR B = 3

EFFECTIVE NUMBER OF OBSERVATIONS FOR ESTIMATION IS    146

INITIAL SUM OF SQUARES =  .1098E+02

ITERATION STOPS - RELATIVE CHANGE IN SUM OF SQUARES LESS THAN  .1000E-02

ITERATION NO.   3

PARAMETER VALUES VIA REGRESSION

       1           2           3
   .4804E+00   .7297E+00   .4709E+01
```

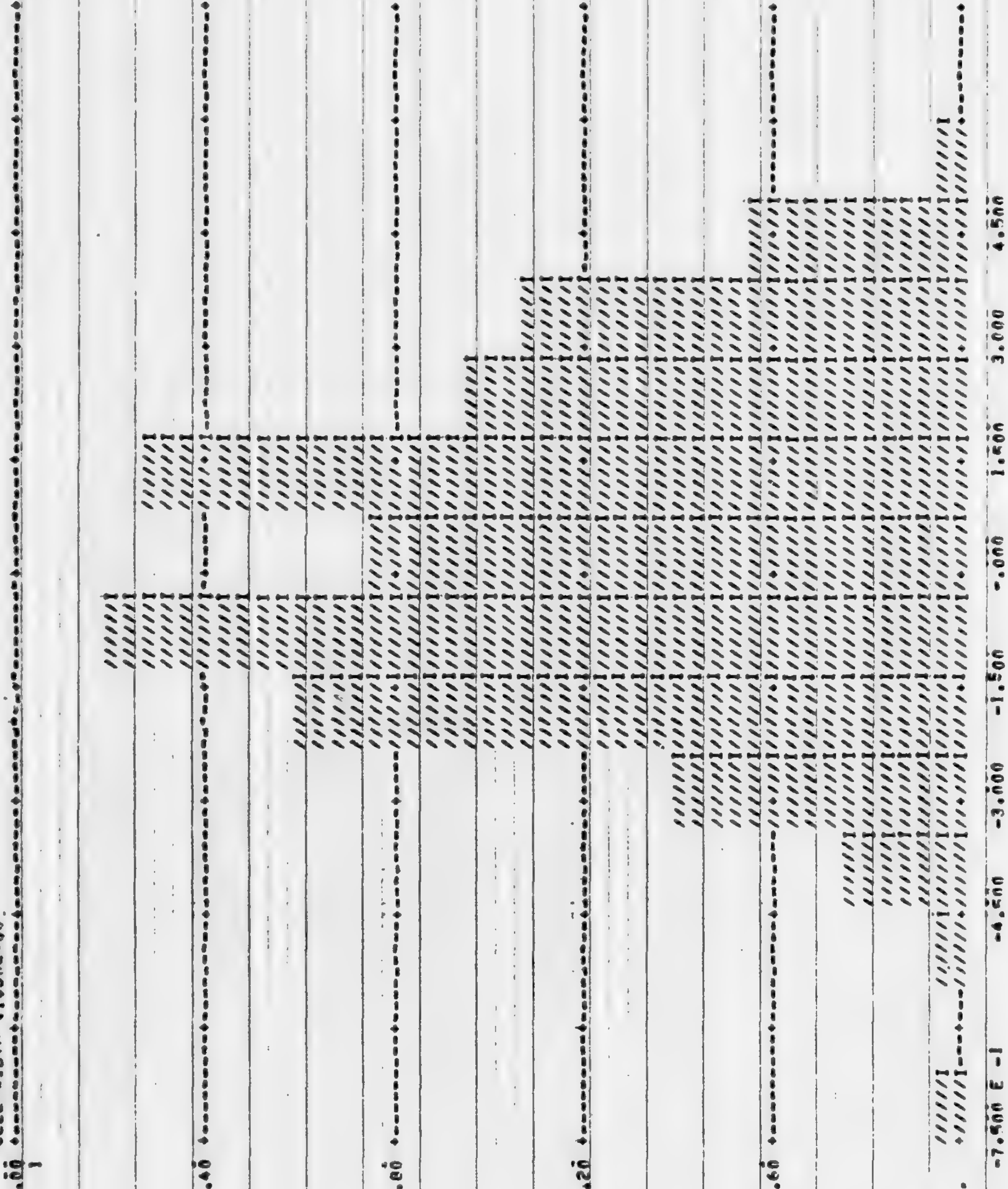
```
SUM OF SQUARES AFTER REGRESSION =  .7420598E+01
************************************************************************************************************************
OPTIMUM VALUE OF B IS  3
************************************************************************************************************************


CORRELATION MATRIX OF THE PARAMETERS
        1            2            3

  1  1.0000

  2  -.0228       1.0000

  3   .0389       -.7111       1.0000

END OF ESTIMATION FOR MODEL  1
```

```
SUMMARY OF MODEL  1
****************************************************************************************************
DATA - X = LEADING INDICATOR                                                    150 OBSERVATIONS
       Y = SALES DATA
DIFFERENCING ON Y - 1 OF ORDER  1
DIFFERENCING ON X - 1 OF ORDER  1
****************************************************************************************************
NOISE MODEL PARAMETERS
****************************************************************************************************
PARAMETER        PARAMETER         PARAMETER      ESTIMATED                95 PER CENT
 NUMBER            TYPE              ORDER          VALUE          LOWER LIMIT          UPPER LIMIT
****************************************************************************************************
   1             MOVING AVERAGE 1      1          .48040E+00        .33384E+00           .62696E+00
****************************************************************************************************
TRANSFER FUNCTION PARAMETERS
****************************************************************************************************
   2             OUTPUT LAG 1          1          .72971E+00        .72158E+00           .73784E+00
   3             INPUT LAG 1           0          .47089E+01        .45934E+01           .48245E+01
****************************************************************************************************
OPTIMUM VALUE OF B IS  3
****************************************************************************************************
OTHER INFORMATION AND RESULTS
****************************************************************************************************
RESIDUAL SUM OF SQUARES     .74286E+01    143 D.F.          RESIDUAL MEAN SQUARE          .51948E-01
NUMBER OF RESIDUALS              146                        RESIDUAL STANDARD ERROR       .22792E+00
```

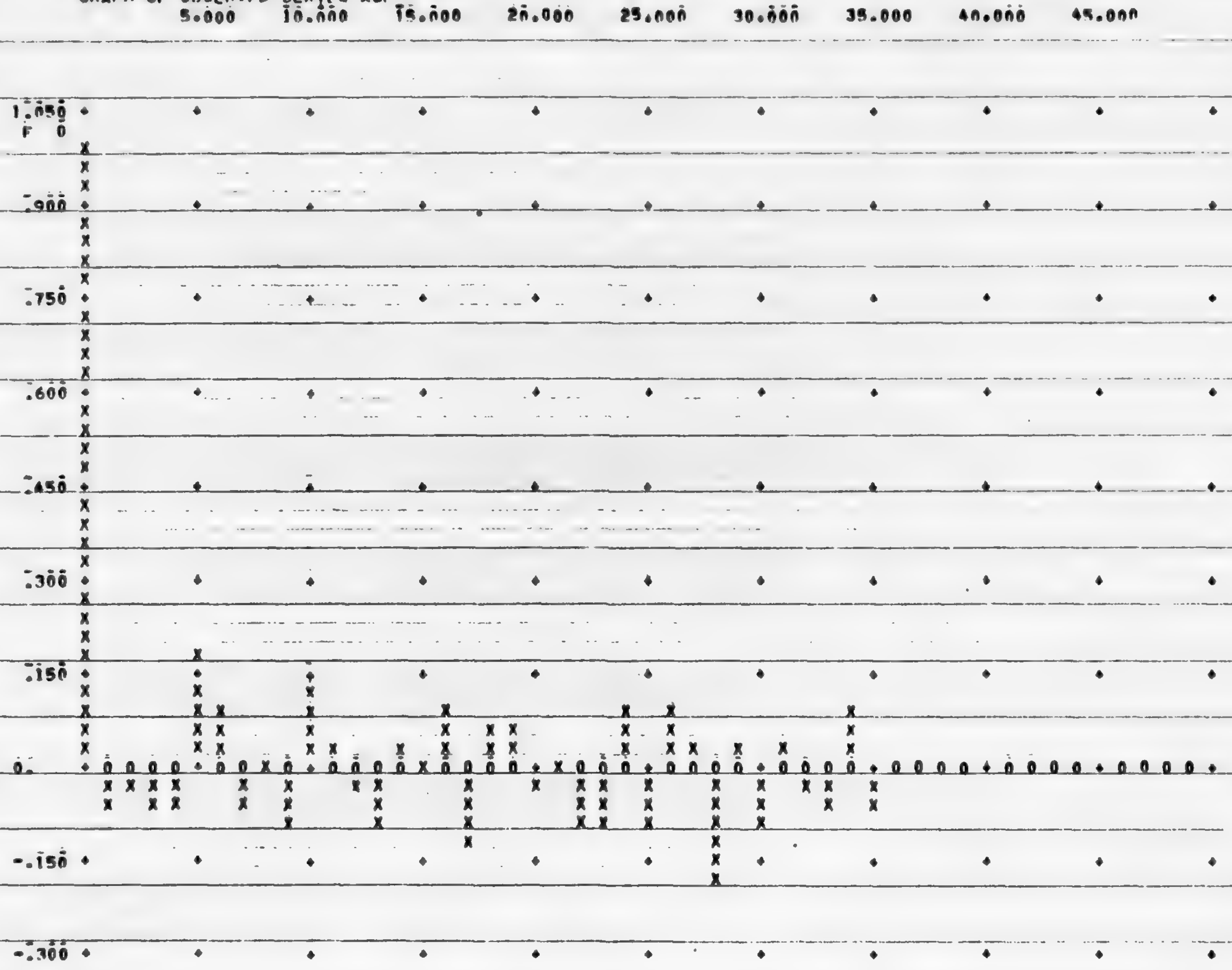
THE ESTIMATED RESIDUALS - MODEL 1
CUMULATIVE PERIODOGRAM .1 PROBABILITY LIMITS



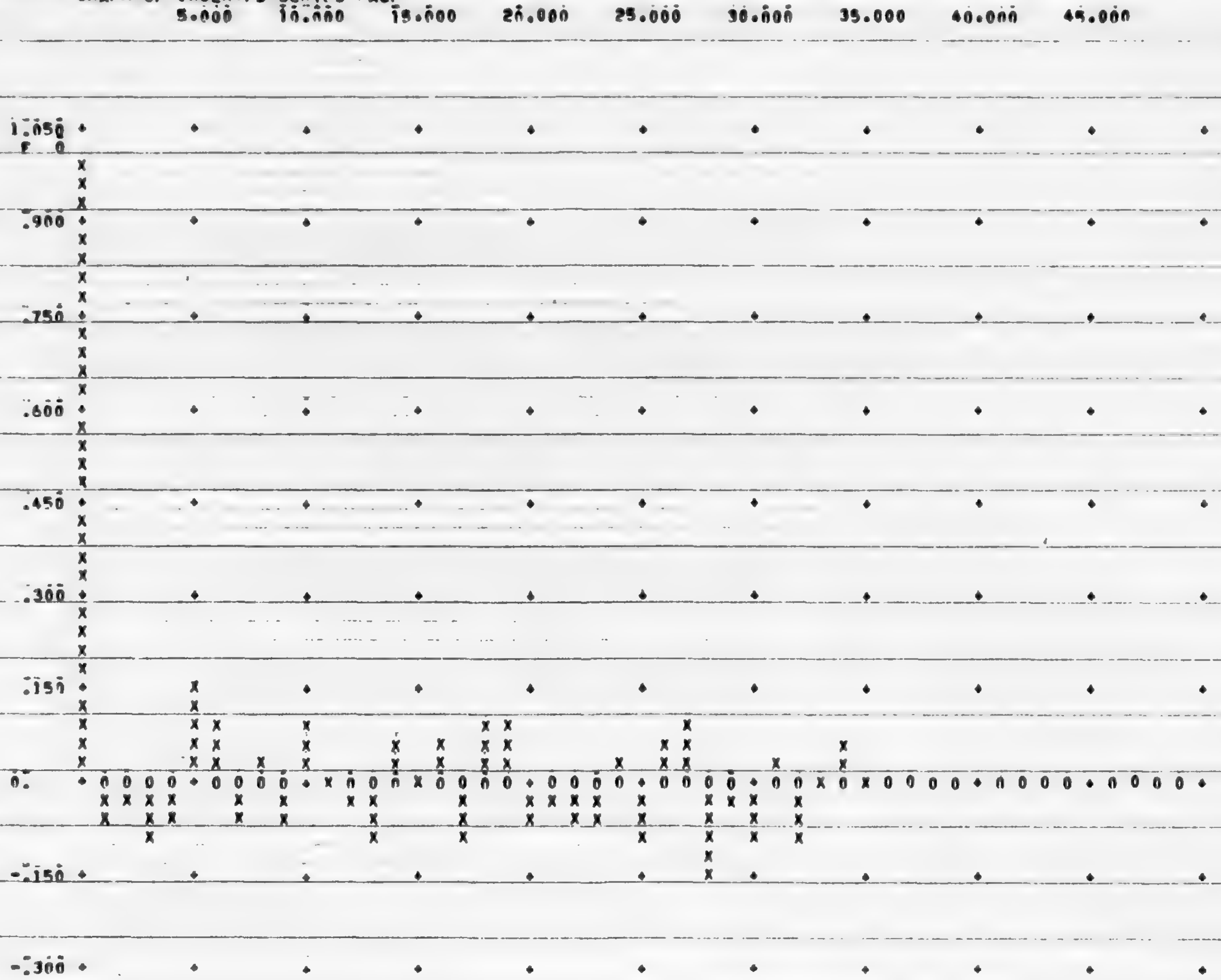
THE ESTIMATED RESIDUALS - MODEL 1
HISTOGRAM
CELL WIDTH .1050E+00



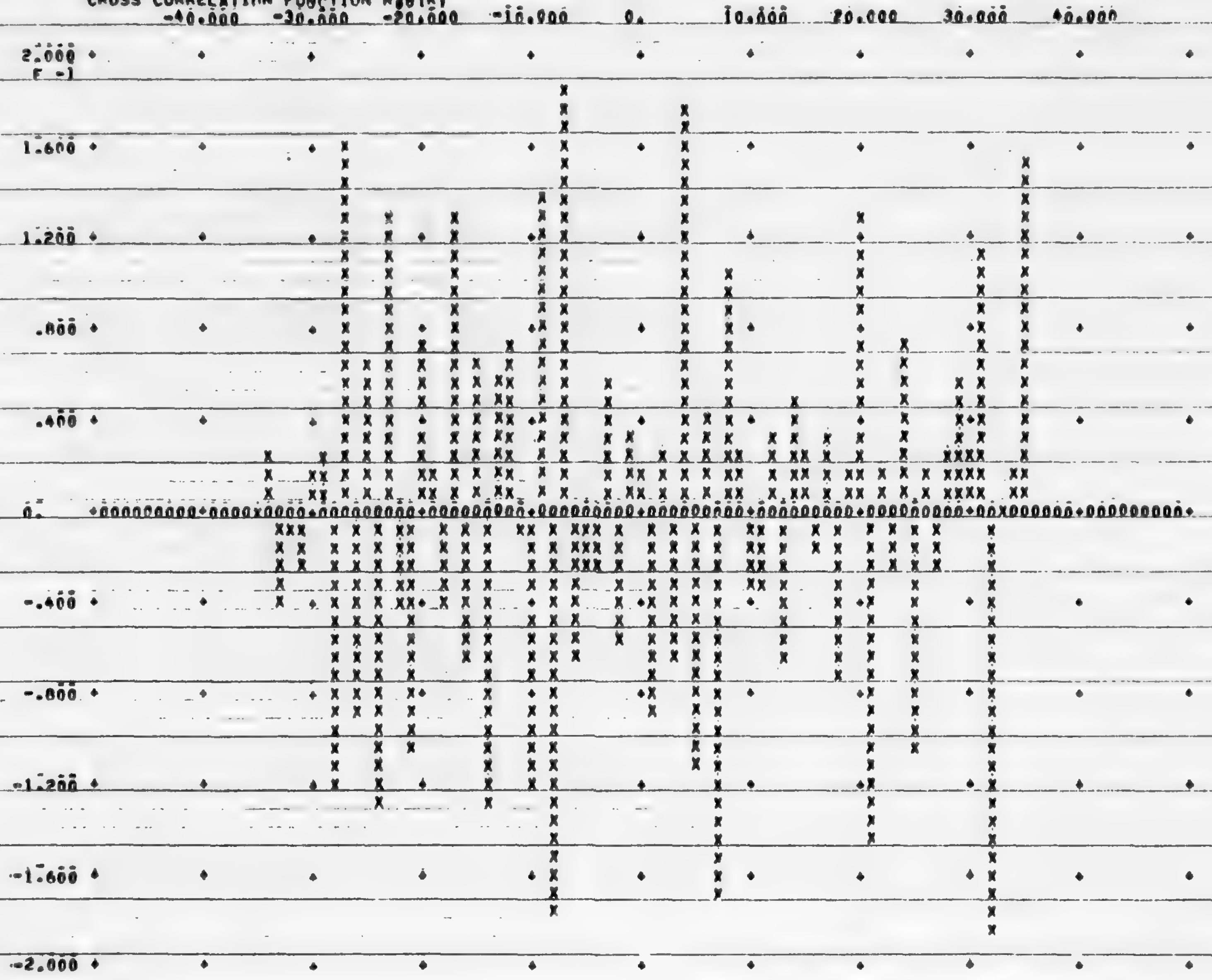
THE ESTIMATED RESIDUALS - MODEL 1
GRAPH OF OBSERVED SERIES ACF



THE ESTIMATED RESIDUALS - MODEL 1
GRAPH OF OBSERVED SERIES PACF



A - PREWHITENED LEADING INDICATOR
B - THE ESTIMATED RESIDUALS - MODEL 1
CROSS CORRELATION FUNCTION RAB(K)

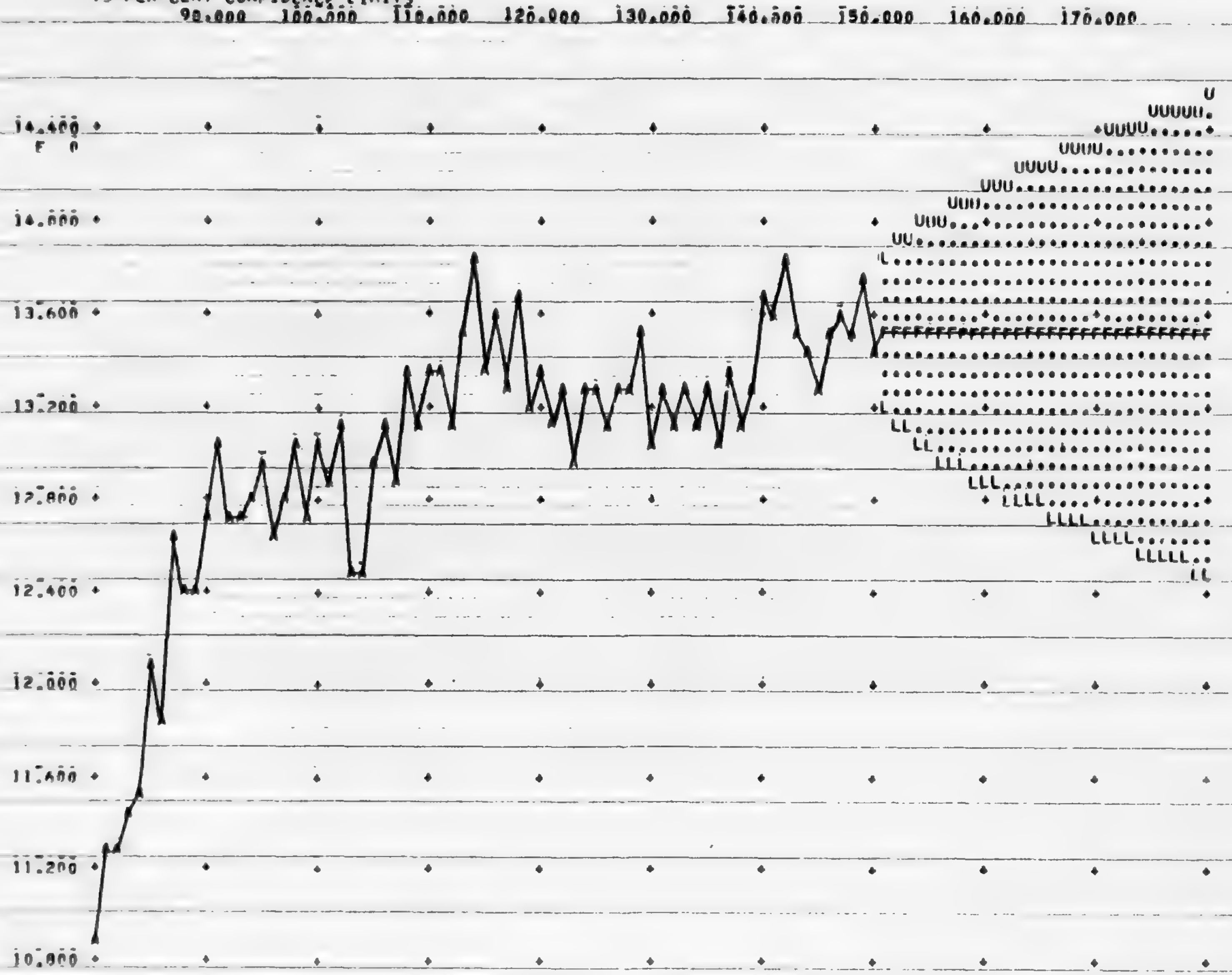


PROGRAM CONTROL STATEMENTS

```
NAME
DATA  150
YNAME
YDATA 150
YRDIF
RDIF
MA    .453,1,1
NMA   .481,1,1
TLAG  4,71,0,1
OLAG  .730,1,1
R     3
TFCST 2,30,150
```

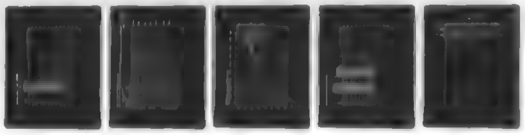
```
TIME SERIES FORECASTING FOR MODEL  1
**********************************************************************************************************************************
DATA - Z = LEADING INDICATOR                                                                                   150 OBSERVATIONS
DIFFERENCING ON Z - 1 OF ORDER  1
**********************************************************************************************************************************
PARAMETER                PARAMETER                PARAMETER          ESTIMATED
 NUMBER                    TYPE                     ORDER              VALUE
**********************************************************************************************************************************
    1                MOVING AVERAGE 1                1               .453300E+00
**********************************************************************************************************************************
NUMBER OF TIME ORIGINS FOR FORECASTS =          1
NUMBER OF FORECASTS AT EACH TIME ORIGIN =      30
FORECAST TIME ORIGINS ARE T=                  150
```

LEADING INDICATOR
FORECAST AT ORIGIN 150   F=FORECAST A=ACTUAL
75 PER CENT CONFIDENCE LIMITS



AD-A035 733 CALIFORNIA UNIV BERKELEY OPERATIONS RESEARCH CENTER F/G 12/1
EVERYMAN'S GUIDE TO TIMES.(U)
JAN 77 R R WILLIE AF-AFOSR-3179-77
UNCLASSIFIED ORC-77-2 NL

2 OF 2
AD A
035 733

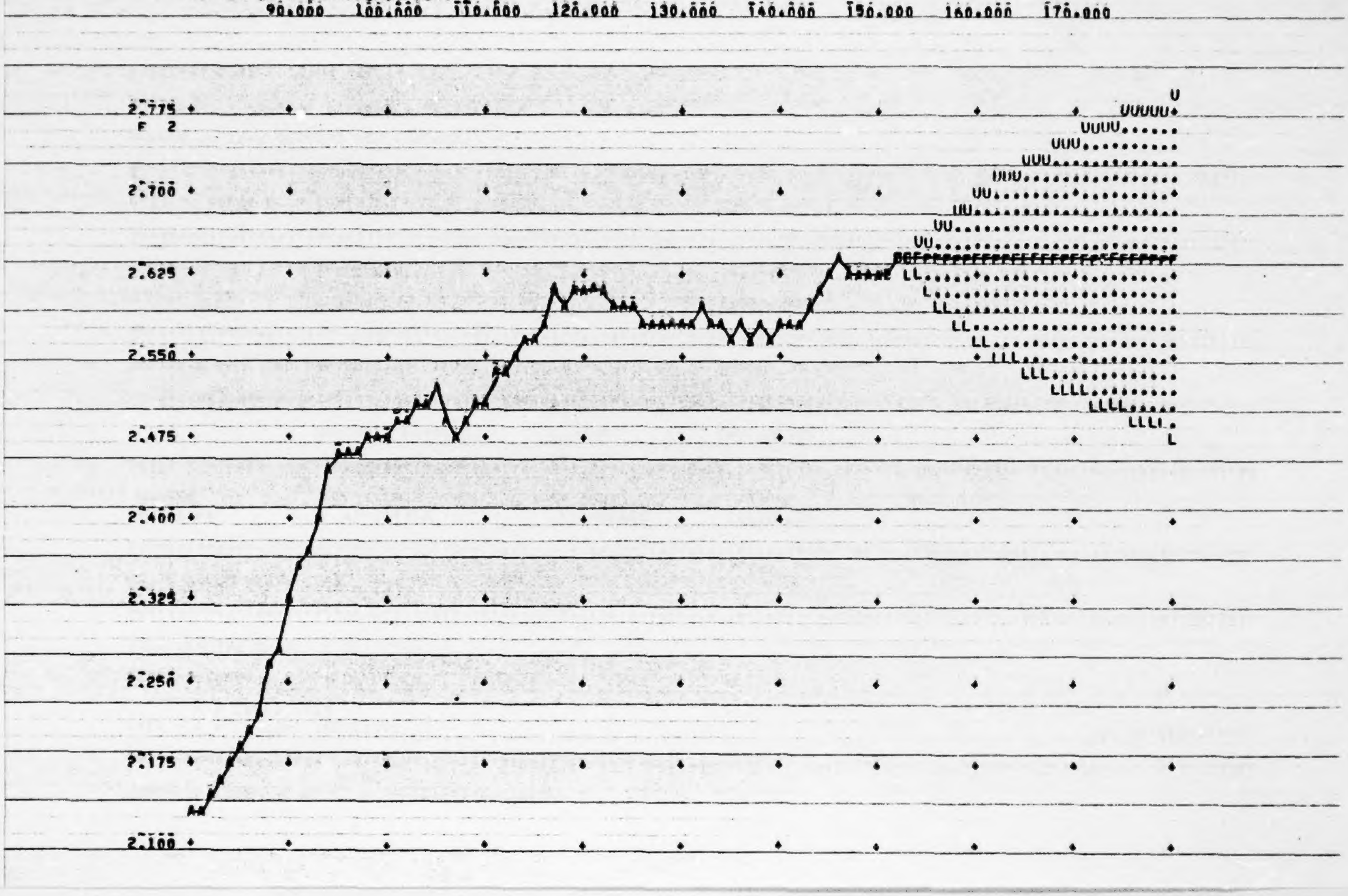


END
DATE
FILMED
3-24-77
NTIS

```
WEIGHTS USED IN CALCULATING CONFIDENCE LIMITS
AND UPDATING FORECASTS AFTER NEW OBSERVATION

      J              PSI(J)

      0           .1000000E+01
      1           .5470000E+00
      2           .5470000E+00
      3           .5470000E+00
      4           .5470000E+00
      5           .5470000E+00
      6           .5470000E+00
      7           .5470000E+00
      8           .5470000E+00
      9           .5470000E+00
     10           .5470000E+00
     11           .5470000E+00
     12           .5470000E+00
     13           .5470000E+00
     14           .5470000E+00
     15           .5470000E+00
     16           .5470000E+00
     17           .5470000E+00
     18           .5470000E+00
     19           .5470000E+00
     20           .5470000E+00
     21           .5470000E+00
     22           .5470000E+00
     23           .5470000E+00
     24           .5470000E+00
     25           .5470000E+00
     26           .5470000E+00
     27           .5470000E+00
     28           .5470000E+00
     29           .5470000E+00
     30           .5470000E+00
```

```
TRANSFER FUNCTION MODEL FORECASTING FOR MODEL 1
****************************************************************************************************
DATA - X = LEADING INDICATOR                                                         150 OBSERVATIONS
       Y = SALES DATA
DIFFERENCING ON Y - 1 OF ORDER  1
DIFFERENCING ON X - 1 OF ORDER  1
****************************************************************************************************
NOISE MODEL PARAMETERS
****************************************************************************************************
PARAMETER           PARAMETER           PARAMETER       ESTIMATED
 NUMBER               TYPE                ORDER           VALUE
****************************************************************************************************
    1             MOVING AVERAGE 1           1          .48100E+00
****************************************************************************************************
TRANSFER FUNCTION PARAMETERS
****************************************************************************************************
    2             OUTPUT LAG 1               1          .73000E+00
    3             INPUT LAG 1                0          .47100E+01
****************************************************************************************************
OPTIMUM VALUE OF B IS  3
****************************************************************************************************
NUMBER OF TIME ORIGINS FOR FORECASTS =       1
NUMBER OF FORECASTS AT EACH TIME ORIGIN =   30
FORECAST TIME ORIGINS ARE T=               150
```

SALES DATA
FORECAST AT ORIGIN 150   F=FORECAST A=ACTUAL
75 PER CENT CONFIDENCE LIMITS



```
WEIGHTS USED IN CALCULATING CONFIDENCE LIMITS

   J               PS(J)             V(J)

   0          .1000000E+01      0.
   1          .5190000E+00      0.
   2          .5190000E+00      0.
   3          .5190000E+00        .4710000E+01
   4          .5190000E+00        .6014670E+01
   5          .5190000E+00        .6967079E+01
   6          .5190000E+00        .7662338E+01
   7          .5190000E+00        .8169877E+01
   8          .5190000E+00        .8540380E+01
   9          .5190000E+00        .8810847E+01
  10          .5190000E+00        .9008289E+01
  11          .5190000E+00        .9152421E+01
  12          .5190000E+00        .9257637E+01
  13          .5190000E+00        .9334445E+01
  14          .5190000E+00        .9390515E+01
  15          .5190000E+00        .9431446E+01
  16          .5190000E+00        .9461325E+01
  17          .5190000E+00        .9483138E+01
  18          .5190000E+00        .9499060E+01
  19          .5190000E+00        .9510684E+01
  20          .5190000E+00        .9519169E+01
  21          .5190000E+00        .9525364E+01
  22          .5190000E+00        .9529885E+01
  23          .5190000E+00        .9533186E+01
  24          .5190000E+00        .9535596E+01
  25          .5190000E+00        .9537355E+01
  26          .5190000E+00        .9538639E+01
  27          .5190000E+00        .9539577E+01
  28          .5190000E+00        .9540261E+01
  29          .5190000E+00        .9540760E+01
  30          .5190000E+00        .9541125E+01
```

REFERENCES

[1] Box, G. E. P. and G. M. Jenkins. TIME SERIES ANALYSIS: FORECASTING AND CONTROL, Holden-Day, (1970).

[2] Brillinger, David R., TIME SERIES: DATA ANALYSIS AND THEORY, Holt, Winston, and Rinehart, (1975).

[3] Jenkins, G. M. and D. G. Watts, SPECTRAL ANALYSIS AND ITS APPLICATIONS, Holden-Day, (1968).